

ANALISIS GEROMBOL KELAS LATEN  
(*LATENT CLASS CLUSTER ANALYSIS*)  
DALAM PENGELOMPOKAN BEBERAPA PRODUK  
PONSEL PINTAR (*SMARTPHONE*)

Skripsi  
Disusun untuk melengkapi syarat-syarat  
guna memperoleh gelar Sarjana Sains



TEDY TRIYADI  
3125111202

PROGRAM STUDI MATEMATIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI JAKARTA  
2015

# LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI

## ANALISIS GEROMBOL KELAS LATEN (*LATENT CLASS CLUSTER ANALYSIS*) DALAM PENGELOMPOKAN BEBERAPA PRODUK PONSEL PINTAR (*SMARTPHONE*)

Nama : Tedy Triyadi

No. Registrasi : 3125111202

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan	: Prof. Dr. Suyono, M.Si. NIP. 19671218 199303 1 005	.....	.....
Wakil Penanggung Jawab			
Pembantu Dekan I	: Dr. Muktiningsih, M.Si. NIP. 19640511 198903 2 001	.....	.....
Ketua	: Ir. Fariani Hermin, M.T. NIP. 19600211 198703 2 001	.....	.....
Sekretaris	: Drs. Bambang Irawan, M.Si. NIP. 19681201 200112 1 001	.....	.....
Penguji	: Drs. Sudarwanto, M.Si., DEA. NIP. 19650325 199303 1 003	.....	.....
Pembimbing I	: Dra. Widyanti Rahayu, M.Si. NIP. 19661103 200112 2 001	.....	.....
Pembimbing II	: Dian Handayani, M.Si. NIP. 19740415 199803 2 001	.....	.....

Dinyatakan lulus ujian skripsi tanggal: 7 Juli 2015

# ABSTRACT

**TEDY TRIYADI, 3125111202. Latent Class Cluster Analysis for Clustering Some Smartphones. Thesis. Faculty of Mathematics and Natural Science Jakarta State University. 2015.**

*The large number of smartphone products which are sold in Indonesia makes customers confused in selecting which product to be purchased. To provide an idea for them regarding the characteristics of smartphone products that are sold in Indonesia, it is necessary to categorize them based on particular characteristics. Therefore, they could obtain information regarding any smartphone possessing similar characteristics. This information will be the one of their considerations in purchasing a smartphone. It also will be costumized with their needs and financial capabilities.*

*Latent Class Cluster Analysis (LCCA) is a clustering method to allocate respondents or objects to several different latent class based on some observed variables which could be numerical or categorical. In this research, parameters model are estimated using maximum likelihood. Meanwhile, to obtain solutions for likelihood equation, we apply iterative Expectation Maximization (EM) algorithm.*

*In clustering 194 types of smartphone as the sample of this research, the result shows there are five smartphone clusters, with the following characteristics: first cluster is the ones with mid-range specification, second cluster is the ones with high-end specification, third cluster is the ones with low-end specification, fourth cluster is the ones with tablet specification, and fifth cluster is the ones with mid-range specification yet having expensive price.*

**Keywords :** *latent class, latent class cluster analysis, maximum likelihood estimation, EM algorithm*

# ABSTRAK

**TEDY TRIYADI, 3125111202. Analisis Gerombol Kelas Laten (*Latent Class Cluster Analysis*) dalam Pengelompokan Beberapa Produk Ponsel Pintar (*Smartphone*). Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. 2015.**

Banyaknya produk *smartphone* dengan jenis yang beragam di pasaran, menjadikan konsumen terkadang bingung dalam memilih produk yang akan dibeli. Untuk memberikan gambaran kepada masyarakat mengenai karakteristik dari beberapa produk *smartphone* yang beredar di Indonesia, maka perlu dilakukan pengelompokan terhadap produk-produk tersebut berdasarkan beberapa karakteristik tertentu, sehingga masyarakat dapat memperoleh informasi produk *smartphone* mana saja yang memiliki kemiripan karakteristik. Informasi ini nantinya akan menjadi dasar pertimbangan untuk membeli produk *smartphone*, yang tentunya juga disesuaikan dengan kebutuhan dan kemampuan finansial.

*Latent Class Cluster Analysis (LCCA)* adalah suatu metode penggerombolan objek atau responden berdasarkan sekumpulan peubah yang teramati ke dalam beberapa gerombol (kelas laten), dimana peubah yang menjadi dasar pengelompokan dapat bertipe campuran antara kategorik dan numerik. Dalam penelitian ini, penaksiran parameter pada model *LCCA* ditaksir dengan menggunakan metode *maximum likelihood*. Sementara itu, untuk mendapatkan solusi dari persamaan likelihood, digunakan algoritma iteratif *Expectation-Maximization (EM)*.

Dalam penggerombolan terhadap 194 tipe *smartphone* yang menjadi sampel dalam penelitian ini, menunjukkan bahwa hasilnya terdapat 5 gerombol *smartphone*, dengan karakteristik sebagai berikut: gerombol pertama adalah gerombol dengan spesifikasi menengah, gerombol kedua adalah gerombol dengan spesifikasi tinggi, gerombol ketiga adalah gerombol dengan spesifikasi rendah, gerombol keempat adalah gerombol dengan spesifikasi *tablet*, dan gerombol kelima adalah gerombol dengan spesifikasi menengah namun memiliki harga paling mahal.

**Kata kunci :** kelas laten, analisis gerombol kelas laten, *maximum likelihood*, algoritma EM.

## PERSEMBAHANKU...

*"So do not weaken and do not grieve, and you will be superior if you are [true] believers." - Q.S. Ali 'Imran [3:139]*

*Before it's called Istanbul, it was called Constantine*

*Before Julius Caesar, there was King Solomon*

*Before Moslem ruled over, Pharaohs ruled Egypt*

*Before Gautama, The Brahmas lectured us*

*And before me?*

*Another one has been conquered, You.*

*In history,*

*It's not about who ruled over*

*Or who conquered*

*Or even who lead you now.*

*In history,*

*It's always about the Greatest*

*It's always about the Unremarkable*

*It's always about the Glorious One*

*And those things that made someone turn to be,*

*Special One. - Anonymous*

Skripsi ini kupersembahkan untuk Bapak, Mama, Adik Kania.

*"every second you live, every breath you take, every smile you give,  
makes me feel alive. I love you".*

## KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas pengetahuan dan kemampuan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Gerombol Kelas Laten (*Latent Class Cluster Analysis*) dalam Pengelompokan Beberapa Produk Ponsel Pintar (*Smartphone*)" yang merupakan salah satu syarat dalam memperoleh gelar Sarjana Sains, Jurusan Matematika Universitas Negeri Jakarta.

Skripsi ini berhasil diselesaikan tidak terlepas dari adanya bantuan dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan terima kasih terutama kepada:

1. Kedua orang tua penulis yang telah merawat, membesarkan, memberikan dukungan, dan doa kepada penulis agar tetap semangat dan pantang menyerah dalam menjalani semua hal sulit yang dihadapi.
2. Ketiga kakak penulis, Ai, Didin, dan Iyan, serta keponakan penulis, Kania yang terus memberi semangat, mendoakan penulis, dan selalu menghibur ketika penulis mengalami kesulitan dalam penulisan skripsi ini.
3. Ibu Dra. Widyanti Rahayu, M.Si. selaku Dosen Pembimbing I dan Ibu Dian Handayani, M.Si. selaku Dosen Pembimbing II, yang telah meluangkan waktunya dalam memberikan bimbingan, saran, nasehat serta arahan sehingga skripsi ini dapat menjadi lebih baik dan terarah. Terima kasih pula karena ibu tidak hanya menjadi pembimbing penulis melainkan sebagai ibu kedua. Mohon maaf atas segala kekurangan, semoga Ibu selalu diberkati Tuhan Yang Maha Esa.

4. Bapak Drs. Makmuri, M.Si., selaku Ketua Jurusan Matematika FMIPA UNJ dan Ibu Ratna Widyati, S.Si, M.Kom., selaku Ketua Prodi Matematika FMIPA UNJ, dan Bapak Drs. Bambang Irawan, M.Si., selaku Pembimbing Akademik atas segala bantuan dan kerja sama Bapak dan Ibu selama pengerjaan skripsi ini.
5. Ibu Ir. Fariani Hermin, M.T., selaku dosen Jurusan Matematika yang sudah mengajari saya beberapa mata kuliah, terutama mata kuliah Statistika, yang kemudian membuat saya cinta akan Statistika, dan Bapak Eka Suryana, M.Kom. yang membantu saya selama pengerjaan skripsi ketika menemui kesulitan, semoga Ibu dan Bapak selalu diberkati Tuhan Yang Maha Esa.
6. Seluruh dosen Matematika FMIPA UNJ yang telah memberikan ilmunya kepada penulis dan seluruh karyawan FMIPA UNJ yang telah memberikan kemudahan kepada penulis dalam bentuk apapun.
7. Teman-teman terbaikku D.O.T Math 2011 Danu, Sandy, Julu, Albert, Bagus, Givatra, Agung, dan yang paling utama Mahdhi, Depe, Hamas, dan Iyus. Terima kasih banyak untuk suka dan duka yang kita hadapi selama ini serta gelak tawa yang kalian berikan.
8. Teman Matematika UNJ 2011. Terutama Ridianti, Danti, Debi, Pephu, Puti, Amchan, Idam, Dytta, Vira, Nida, Monik, Riska, Syifa, Gia, Firda, Tyan, dll. Terima kasih atas perjuangannya selama ini.
9. Sahabatku Siti Yulia dan Murni Susilo Yekti, yang selalu memberikan kebahagiaan dan hiburan. Terima kasih atas dukungan dan doa yang kalian berikan selama ini.

10. Sahabat-sahabat terbaikku dari OSIS 22: Sebastianus, Desy, Sisilia, Rangga, Pratiwi, Raisya, Oky, Diaz, dan Jusup. Terima kasih atas kebersamaan yang sudah kita bangun selama ini, terima kasih pula atas doa dan dukungan kalian.
11. Sahabat-sahabat terbaikku dari PB: Avior, Ifnu, Randy, Johannes, dan Ikhwan. Terima kasih atas kebersamaan, doa, dan dukungan kalian selama ini.
12. Teman-teman DeuxTrois 2011, yang memberikan masa-masa indah. Terima kasih atas doa, semangat, dukungan, suka dan duka dari kalian. Sukses untuk kita semua.
13. Kakak-kakakku, Kak Ria, Kak Fitri, Kak Gusti, Kak Eche, dan Kak Shinta. Terima kasih atas bantuan dan dukungannya selama ini.
14. Murid-muridku yang sudah seperti adik-adikku sendiri: Wahyu, Hellena, Fathan, Ilham, Prilly, Claudia, Gebi, Michel, Mahesa, Hestin, Zahra, Ateng, Billy, Sulas, Ditri, Darell, dll. Terima kasih atas semangat, doa, dan senyuman yang kalian berikan setiap hari.
15. Terakhir untuk Fenny Sarfina dan Richard Sitorus. Terima kasih sudah menjadi Teman Hidup dan Gajah untuk penulis, terima kasih pula atas dukungan, doa, dan kebaikan-kebaikan kalian selama ini, semoga segala kebaikan kalian dibalas Tuhan Yang Maha Esa.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Masukan dan kritikan akan sangat berarti. Semoga skripsi ini dapat bermanfaat bagi pembaca sekalian.

Jakarta, Juli 2015

Tedy Triyadi

# DAFTAR ISI

<b>ABSTRACT</b>	<b>i</b>
<b>ABSTRAK</b>	<b>ii</b>
<b>KATA PENGANTAR</b>	<b>iv</b>
<b>DAFTAR ISI</b>	<b>x</b>
<b>DAFTAR TABEL</b>	<b>xii</b>
<b>DAFTAR GAMBAR</b>	<b>xiii</b>
<b>I PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang Masalah . . . . .	1
1.2 Perumusan Masalah . . . . .	4
1.3 Pembatasan Masalah . . . . .	4
1.4 Tujuan Penulisan . . . . .	5
1.5 Manfaat Penulisan . . . . .	5
1.6 Metode Penelitian . . . . .	5
<b>II LANDASAN TEORI</b>	<b>6</b>
2.1 Analisis Gerombol . . . . .	6
2.2 Jenis Skala Data . . . . .	9
2.3 Keluarga Eksponensial . . . . .	11
2.3.1 Distribusi <i>Bernoulli</i> . . . . .	12
2.3.2 Distribusi Multinomial . . . . .	14

2.3.3	Distribusi Normal . . . . .	15
2.4	Teorema Bayes . . . . .	16
2.5	Uji Normalitas . . . . .	17
2.6	<i>Latent Class Cluster Analysis</i> . . . . .	19
2.6.1	Model <i>Latent Class Cluster Analysis</i> untuk Peubah Kategorik	22
2.6.2	Model <i>Latent Class Cluster</i> untuk Peubah Kontinu . . . . .	26
2.7	Penaksiran Maksimum <i>Likelihood</i> . . . . .	27
2.8	Algoritma EM ( <i>Expectation-Maximization</i> ) . . . . .	29
2.9	Pemilihan Model Terbaik . . . . .	31
2.9.1	AIC ( <i>Akaike's Information Criterion</i> ) . . . . .	31
2.9.2	BIC ( <i>Bayesian Information Criterion</i> ) . . . . .	31
2.10	Uji Kelayakan Model . . . . .	32
2.11	Uji Asumsi Kebebasan Lokal . . . . .	33
2.12	Klasifikasi Objek . . . . .	34
2.13	Keragaman Gerombol . . . . .	34
2.14	Karakteristik <i>Smartphone</i> . . . . .	37
<b>III</b>	<b>PEMBAHASAN</b>	<b>42</b>
3.1	Model <i>Latent Class Cluster Analysis</i> Berdasarkan Peubah Bertipe Campuran . . . . .	42
3.2	Penaksiran Parameter . . . . .	45
3.3	Pemilihan Model . . . . .	49
3.4	Data . . . . .	50
3.4.1	Sumber Data . . . . .	50
3.4.2	Identifikasi Peubah . . . . .	51
3.5	Metode Analisis Data . . . . .	54

3.6 Hasil Analisis Data . . . . .	58
3.6.1 Penaksiran Peluang Awal Kelas Laten . . . . .	58
3.6.2 Pemilihan Model Terbaik . . . . .	59
3.6.3 Uji Asumsi Kebebasan Lokal Model Terpilih . . . . .	61
3.6.4 Taksiran Parameter Model Terpilih . . . . .	64
3.6.5 Pengklasifikasian Objek Pada Setiap Gerombol . . . . .	71
3.6.6 Rasio Keragaman Gerombol Model Terpilih . . . . .	83
<b>IV PENUTUP</b>	<b>85</b>
4.1 Kesimpulan . . . . .	85
4.2 Saran . . . . .	86
<b>DAFTAR PUSTAKA</b>	<b>87</b>
<b>LAMPIRAN-LAMPIRAN</b>	<b>91</b>

## DAFTAR TABEL

2.1	Perbedaan <i>K-Means</i> dengan <i>LCCA</i> . . . . .	22
3.1	Contoh Perhitungan BIC . . . . .	50
3.2	Peubah Teramati yang Digunakan . . . . .	51
3.3	Peubah Teramati yang Digunakan (Lanjutan) . . . . .	52
3.4	Uji Normalitas Data Sebelum Ditransformasi . . . . .	53
3.5	Uji Normalitas Data Setelah Ditransformasi . . . . .	53
3.6	Peluang Awal Kelas Laten ( $\eta_j$ ) Untuk Setiap Kemungkinan Model	58
3.7	Kriteria Untuk Pemilihan Model Terbaik . . . . .	59
3.8	Alokasi Objek Pada Setiap Gerombol . . . . .	61
3.9	Kriteria Untuk Pemilihan Model Terbaik dengan Menggunakan <i>Direct Effect</i> . . . . .	62
3.10	Alokasi Objek Pada Setiap Gerombol Untuk Model dengan Enam <i>Direct Effect</i> . . . . .	64
3.11	Taksiran Peluang Masing-Masing Kategori Pada Peubah Biner . . . . .	65
3.12	Taksiran Peluang Masing-masing Kategori Pada Peubah Nominal . . . . .	66
3.13	Taksiran Peluang Masing-masing Kategori Pada Peubah Ordinal . . . . .	68
3.14	Taksiran <i>Mean</i> dan <i>Varians</i> Pada Peubah Kontinu . . . . .	70
3.15	Alokasi Produk <i>Smartphone</i> Pada Gerombol Pertama . . . . .	73
3.16	Alokasi Produk <i>Smartphone</i> Pada Gerombol Kedua . . . . .	75
3.17	Alokasi Produk <i>Smartphone</i> Pada Gerombol Ketiga . . . . .	78
3.18	Alokasi Produk <i>Smartphone</i> Pada Gerombol Keempat . . . . .	80
3.19	Alokasi Produk <i>Smartphone</i> Pada Gerombol Kelima . . . . .	81

3.20 Rasio Keragaman Pada Peubah Kontinu . . . . .	83
3.21 Rasio Keragaman Pada Peubah Kategorik . . . . .	84

## DAFTAR GAMBAR

3.1	Diagram Alir <i>Latent Class Cluster Analysis</i> . . . . .	54
3.2	Diagram Alir <i>Latent Class Cluster Analysis</i> (Lanjutan) . . . . .	55
3.3	Diagram Alir Algoritma EM . . . . .	56
3.4	Nilai BIC Model 1 Gerombol Hingga 5 Gerombol . . . . .	60
3.5	Nilai BIC dengan Penambahan <i>Direct Effect</i> . . . . .	63

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Perkembangan teknologi informasi dan komunikasi saat ini sangat pesat, begitu pula inovasi yang dilakukan oleh para ahli di bidangnya. Salah satu contoh perkembangan teknologi informasi yang meningkat pesat dalam kurun waktu belakangan ini adalah produk *gadget* yang dikenal dengan nama ponsel pintar (*smartphone*), dalam bentuk *tablet* maupun *phablet*. Saat ini produsen *smartphone* berlomba-lomba merebut pangsa pasar teknologi informasi dan komunikasi yang tidak ada habisnya.

Menurut Yang *et al* (2007), ponsel pintar (*smartphone*) merupakan telepon seluler yang menggabungkan fungsi-fungsi *Personal Digital Assistant* (PDA) seperti kalender, *personal schedule*, *address book*, dan memiliki kemampuan untuk mengakses internet, membuka email, membuat dokumen, bermain game, serta membuka aplikasi lainnya. Dengan adanya produk *smartphone* ini diharapkan masyarakat semakin dipermudah dalam mobilitas sehari-hari, meskipun kemunculan *smartphone* terkadang membawa pengaruh buruk bagi penggunanya itu sendiri terlebih jika diberikan kepada anak dibawah umur yang harus memerlukan pengawasan dalam mengakses informasi. Semakin tingginya tingkat antusias masyarakat terhadap perkembangan teknologi, menyebabkan para produsen *smartphone* ini mengeluarkan banyak produk yang beragam baik dilihat dari spe-

sifikasi komponen internal *smartphone* itu sendiri, bentuk fisik, sistem operasi yang digunakan, fitur-fitur penunjang (kamera, *flash*, pemutar musik, dsb), harga, ukuran layar, maupun hal-hal yang berhubungan dengan *smartphone* itu sendiri.

Analisis gerombol (*cluster analysis*) adalah suatu analisis peubah ganda yang bertujuan untuk mensegmentasikan amatan ke dalam suatu gerombol dengan variasi objek dalam satu gerombol yang sama adalah sekecil mungkin, sedangkan variasi antar gerombol semaksimal mungkin. Analisis gerombol bermanfaat untuk mengelompokkan objek berdasarkan kesamaan karakteristik di antara objek-objek tersebut. Dengan demikian, ciri-ciri suatu gerombol yang baik yaitu mempunyai homogenitas internal (*within cluster*) yaitu kesamaan antar anggota dalam satu gerombol serta homogenitas eksternal (*between cluster*) yaitu perbedaan antara gerombol yang satu dengan gerombol yang lain (Handayani, 2010).

Permasalahan yang biasa ditemukan dalam analisis gerombol adalah mengenai penentuan gerombol yang akan dibentuk dan bagaimana melakukan penggerombolan terhadap data yang berukuran besar dengan tipe data campuran, yaitu data numerik dan kategorik baik dikotomus maupun politomus. Data dikotomus adalah data kategorik yang hanya terdiri dari dua pilihan, sedangkan data politomus adalah data kategorik yang memiliki pilihan lebih dari dua. Menurut Umah (2014), adapun metode yang dapat digunakan untuk proses penggerombolan data campuran diantaranya Analisis Gerombol Hibrid (*Hybrid Clustering*), Analisis Gerombol Dua Tahap (*Two Step Cluster Analysis*), maupun Analisis Gerombol Kelas Laten (*Latent Class Cluster Analysis*).

Analisis gerombol kelas laten (*Latent Class Cluster Analysis*) yang diperkenalkan oleh Lazarsfeld pada tahun 1963 merupakan metode penggerombolan yang digunakan untuk mengidentifikasi keanggotaan kelompok yang tidak teramati (*unobservable*) antara objek dengan peubah yang diamati dengan proses

penggerombolan yang tidak jauh berbeda dengan analisis gerombol *K-Means*. Kedua metode tersebut mengelompokkan sejumlah objek yang terdekat dengan pusatnya. Jika analisis gerombol *K-Means* menggunakan konsep jarak sebagai dasar penggerombolan yang biasanya hanya digunakan untuk data kontinu saja, maka *LCCA* menggunakan konsep peluang bersama sebagai dasar penggerombolan, sehingga dapat digunakan untuk data campuran. Parameter-parameter dalam metode *LCCA* ini ditaksir melalui metode *Maximum Likelihood* (Vermunt dan Magidson, 2002).

Terdapat beberapa penelitian sebelumnya terkait *LCCA* ini yaitu Hidayat (2009) membandingkan metode dari *LCCA* dan *K-Means* dalam mengelompokkan sekelompok data kontinu dalam kasus pengelompokkan kabupaten/kota di wilayah Jawa Barat dan Bali, hasilnya metode *LCCA* lebih optimal daripada *K-Means*, lalu Maulidya (2009) menerapkan *LCCA* untuk segmentasi pasar terhadap peubah nominal dan kontinu yang diamati dalam studi kasus pengelompokkan konsumen *K-Liquid Chlorophyll* di wilayah kota Malang. Selanjutnya, Handayani (2010) menjelaskan tentang bagaimana penaksiran parameter *Latent Class Model* dengan menggunakan algoritma EM pada peubah biner, dan Umah (2014) membandingkan hasil pengklasifikasian antara *LCCA* dengan *Two Step Cluster Analysis* terhadap beberapa studi kasus, hasilnya pengelompokkan dengan menggunakan *LCCA* lebih optimal daripada *Two Step Cluster Analysis*. Kemudian pada tahun 2014, Rendra mengelompokkan sejumlah pengguna *Flexy Trendy* di Jawa Timur menggunakan *LCCA* berdasarkan peubah ordinal yang diamati, dan Budiati (2014) melakukan pengelompokkan daerah penghasil bahan dasar tepung komposit di Indonesia dengan *LCCA* berdasarkan peubah kontinu dan nominal yang diamati.

Dalam skripsi ini akan dibahas pengelompokan beberapa produk *smartphone* dari berbagai produsen berdasarkan peubah/kriteria seperti jenis sistem operasi, kamera, kapasitas *memory*, ukuran fisik, daya tahan baterai, dan sebagainya. Peubah-peubah tersebut bertipe kategorik dan numerik.

## 1.2 Perumusan Masalah

Perumusan masalah yang akan dikaji adalah sebagai berikut:

1. Bagaimana penaksiran parameter model *LCCA*?
2. Bagaimana karakteristik produk *smartphone* dari berbagai produsen yang sedang beredar di masyarakat?.
3. Bagaimana karakteristik dari setiap gerombol produk *smartphone* yang terbentuk berdasarkan *LCCA*?

## 1.3 Pembatasan Masalah

Pembatasan masalah dalam penulisan ini adalah:

1. Peubah yang mendasari pengelompokan *smartphone* memiliki tipe numerik dan kategorik. Untuk peubah dengan tipe numerik diasumsikan berdistribusi normal. Untuk peubah kategorik diasumsikan berdistribusi *Bernoulli* dan multinomial.
2. Penaksiran parameter model *LCCA* dilakukan secara iterasi numerik dengan algoritma EM.

## 1.4 Tujuan Penulisan

Tujuan yang ingin dicapai dalam skripsi ini adalah mengetahui karakteristik pengelompokan (*clustering*) produk-produk *smartphone* dengan menggunakan metode analisis gerombol kelas laten (*LCCA*).

## 1.5 Manfaat Penulisan

Manfaat yang diharapkan dari skripsi ini adalah menambah wawasan kepada masyarakat/pengguna *smartphone* tentang pengelompokan produk *smartphone* beserta karakteristiknya sehingga dapat membantu konsumen dalam memilih produk.

## 1.6 Metode Penelitian

Skripsi ini merupakan kajian dalam bidang statistika yang didasarkan pada buku-buku dan jurnal-jurnal tentang teori permasalahan di bidang analisis peubah ganda. Referensi utama yang digunakan adalah Moustaki dan Papageorgiou (2004).

## BAB II

# LANDASAN TEORI

### 2.1 Analisis Gerombol

Analisis gerombol merupakan suatu teknik yang digunakan untuk mengklasifikasikan objek atau kasus (responden) ke dalam kelompok sedemikian sehingga objek dalam satu kelompok yang sama relatif homogen, yang dinamakan gerombol. Objek dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh dengan objek dari gerombol lainnya. Di dalam penggerombolan setiap objek hanya boleh masuk ke dalam satu gerombol saja sehingga tidak terjadi tumpang tindih (*overlapping* atau *interaction*). Analisis gerombol mengelompokkan objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam gerombol yang sama. Gerombol-gerombol yang terbentuk memiliki homogenitas internal dan eksternal yang tinggi. Homogenitas internal (*within cluster*) yaitu kesamaan antar anggota dalam satu gerombol serta homogenitas eksternal (*between cluster*) yaitu perbedaan antara gerombol yang satu dengan gerombol yang lain.

Tujuan dari analisis gerombol sendiri adalah mengeksplorasi dan membentuk suatu klasifikasi, menyederhanakan data (*summarizing data*), serta mengidentifikasi hubungan antara objek atau peubah (*relationship identification*). Pembentukan kelompok atau gerombol dalam analisis ini berdasarkan jarak, diharapkan dengan terbentuknya gerombol tersebut lebih mudah menganalisis dan lebih

tepat dalam pengambilan keputusan yang berkaitan dengan masalah tersebut.

Secara umum terdapat dua metode pengelompokan data yaitu metode pengelompokan hierarki dan non hierarki. Metode pengelompokan hierarki adalah metode pengelompokan data yang mengelompokkan  $n$  data ke dalam  $k$  kelompok yang belum ditentukan sebelumnya, sedangkan metode pengelompokan non hierarki adalah metode pengelompokan yang mengelompokkan  $n$  data ke dalam  $k$  kelompok yang sudah ditentukan terlebih dahulu.

Metode hierarki efektif digunakan untuk melakukan penggerombolan terhadap data yang berukuran tidak terlalu besar. Metode hierarki digunakan apabila banyaknya gerombol yang akan dibentuk belum diketahui sebelumnya. Terdapat dua prosedur yang termasuk dalam metode hierarki, yaitu penggabungan (agglomerasi) dan pembagian (divisi).

Dalam metode agglomerasi tiap observasi pada mulanya dianggap sebagai gerombol tersendiri sehingga terdapat gerombol sebanyak jumlah observasi. Kemudian dua gerombol yang terdekat kesamaanya digabung menjadi suatu gerombol baru, sehingga jumlah gerombol berkurang satu pada setiap tahap. Sebaliknya pada pemecahan (*divisive*) dimulai dari satu gerombol besar yang mengandung seluruh observasi, selanjutnya observasi-observasi yang paling tidak sama dipisah dan dibentuk gerombol-gerombol yang lebih kecil. Proses ini dilakukan hingga tiap observasi menjadi gerombol tersendiri. Metode *divisive* memulai pengelompokan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian, proses pengelompokan diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya, sehingga gerombol akan membentuk semacam pohon atau *dendogram* di mana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip.

Dalam metode non hierarki ditentukan terlebih dahulu jumlah gerombol yang diinginkan pada akhir penggerombolan. Salah satu contoh metode penggerombolan non hierarki adalah metode *K-Means*. Berbeda dengan metode hierarki, pusat gerombol yang dipilih pada metode ini merupakan pusat gerombol sementara (*interim centers*) dan pusat gerombol ini terus diperbaharui pada tiap iterasi sampai kriteria pemberhentian tercapai. Dengan demikian objek yang telah berada pada suatu gerombol tertentu dimungkinkan tidak dapat pindah ke gerombol yang lain.

Metode *K-Means* merupakan salah satu metode analisis gerombol non hierarki yang digolongkan sebagai metode penggerombolan yang bersifat *unsupervised* (tanpa arahan) karena data yang dianalisis tidak mempunyai label kelas, yang berarti dalam proses penggerombolannya tidak mempunyai anggota gerombol yang pasti. Metode *K-Means* ini disebut juga sebagai metode *semi-supervised classification*. karena para peneliti sering menentukan sendiri jumlah gerombol awal sesuai kebutuhan.

Istilah *K-Means* sendiri ditemukan untuk mendeskripsikan bahwa algoritma ini menandai setiap objek masuk ke dalam kelompok (gerombol) yang mempunyai rata-rata (pusat gerombol) terdekat. Langkah-langkah dalam analisis gerombol *K-Means* dapat disusun dengan tahapan sebagai berikut:

1. Menentukan jumlah gerombol,
2. Menghitung jarak antara objek dengan tiap *centroid*,
3. Menghitung kembali rata-rata (*centroid*) untuk gerombol yang baru terbentuk,

4. Mengulangi langkah (2) dan (3) sampai tidak ada lagi perpindahan objek antar gerombol (Sartono, 2003).

Terdapat beberapa hal yang dapat mempengaruhi hasil penggerombolan antara lain:

1. Pencilan (*outlier*)

Pencilan (*outlier*) adalah suatu objek yang berbeda atau sangat terpisah dengan objek lainnya. Jika dalam proses penggerombolan terdapat pencilan maka cara yang sering dilakukan adalah dengan tidak mengikutsertakan pencilan dalam hasil akhir penggerombolan atau memisahkan pencilan menjadi sebuah kelompok tersendiri.

2. Kolinearitas diantara dua Peubah

Kolinearitas (*collinearity*) dapat didefinisikan sebagai korelasi (*correlation*) antara dua peubah yang terlibat dalam suatu multikolinearitas. Masalah multikolinearitas dalam penggerombolan disebabkan oleh adanya peubah yang tidak saling bebas sehingga peubah tersebut perlu ditangani sebelum memulai proses penggerombolan. Jika di antara dua peubah terdapat korelasi yang tinggi maka pengurangan (*elimination*) peubah yang tidak saling bebas (*dependent*) merupakan cara untuk mengatasinya (Sambandam, 2010).

## 2.2 Jenis Skala Data

Menurut Sukestiyarno (2008), jenis data berdasarkan skala pengukuran, dibagi menjadi::

## 1. Data Nominal

Data ini berbentuk bilangan diskrit dan merupakan hasil konversi data kualitatif. Tiap bilangan dari jenis data ini tidak mempunyai arti menurut besarnya ataupun posisinya, melainkan hanya sebagai simbolisasi data saja. Data dapat secara bebas disusun tanpa memperhatikan urutan, dan dapat dipertukarkan sesuai kesepakatan. Contoh:

- Peubah jenis agama. Islam = 1, Kristen = 2, Katolik = 3, Hindu = 4, Buddha = 5, dan Konghucu = 6.
- Peubah status pernikahan. Lajang = 1, Sudah Menikah = 2, Bercerai = 3.
- Peubah jenis kelamin. Laki-laki = 1, Perempuan = 2.

Bilangan-bilangan tersebut dapat dipertukarkan sesuai kesepakatan tidak akan mempengaruhi skalanya.

## 2. Data Ordinal

Data ini berasal dari hasil pengamatan, observasi, atau angket dari suatu pengamatan. Seperti data nominal, data ini juga merupakan hasil konversi dari data kualitatif, di mana bilangan konversinya menunjukkan urutan menurut kualitas atributnya. Contoh :

- Peubah kinerja mahasiswa: 1 = Sangat Jelek, 2 = Jelek, 3 = Cukup, 4 = Bagus, 5 = Sangat Bagus.
- Peubah kepuasan konsumen. 1 = Sangat Puas, 2 = Kurang Puas, 3 = Cukup Puas, 4 = Sangat Puas.

Bilangan pengganti kualitas tersebut mempunyai suatu tingkatan atribut.

### 3. Data Interval

Data ini merupakan hasil dari pengukuran suatu peubah. Data interval merupakan data kontinu yang mempunyai urutan. Pada data jenis ini tidak mempunyai nol mutlak. Artinya jika responden mempunyai peubah bernilai nol (0) bukan berarti tidak memiliki substansi sama sekali. Misalkan pada peubah suhu/temperatur suatu ruangan terdapat ruangan yang mempunyai suhu  $0^{\circ}C$ , disini bukan berarti ruangan tersebut tidak ada temperatur sama sekali tetapi suhu  $0^{\circ}C$  masih bermakna mempunyai substansi suhu, bahkan masih ada suhu negatif.

### 4. Data Rasio

Data ini berasal dari hasil mengukur suatu peubah. Data ini diasumsikan berbentuk bilangan kontinu hampir sama dengan skala interval, perbedaannya terletak pada nilai nol (0). Pada data rasio mempunyai nol (0) mutlak, artinya jika suatu responden peubahnya bernilai nol (0) berarti tidak memiliki substansi sama sekali. Misalkan peubah massa benda, bila berbicara suatu benda massanya nol (0) kg berarti benda itu tidak ada barangnya.

## 2.3 Keluarga Eksponensial

Keluarga eksponensial merupakan sebuah himpunan yang berisikan distribusi yang berasal dari peubah acak diskrit dan kontinu. Sifat-sifat dalam setiap distribusi ini dapat digeneralisasikan dalam bentuk yang sama. Beberapa

distribusi peluang yang termasuk dalam kategori ini adalah distribusi-distribusi yang kemudian digunakan untuk setiap skala data yang digunakan.

**Definisi 2.3.1.** Fungsi kepadatan peluang untuk peubah random saling bebas dan berdistribusi  $f(x; \theta)$  dengan  $\theta \in \Theta \subseteq \mathbb{R}$  dengan  $\Theta$  adalah ruang parameter, dinamakan keluarga eksponensial jika fungsi tersebut dapat dibentuk menjadi

$$f(x; \theta) = C(\theta)h(x) \exp[Q(\theta)T(x)] \quad (2.1)$$

dengan  $x \in \mathbb{R}$ ,  $\theta \in \Theta$ ,  $C(\theta) > 0$ , dan  $h(x) > 0$  untuk setiap  $x \in S$ . Dimana  $Q(\theta)$  (atau sering disebut  $\theta^T$ ) adalah fungsi vektor parameter dari  $\theta$ ,  $T(x)$  adalah vektor statistik cukup dari fungsi distribusi tersebut,  $h(x)$  adalah fungsi untuk peubah acak  $x$ , biasanya berbentuk konstanta, dan  $C(\theta)$  adalah *cumulative generating function* yang memastikan fungsi distribusi terintegralkan menuju 1. Dimana:

$$C(\theta) = \int h(x) \exp[Q(\theta)T(x)] dx = 1 \quad (2.2)$$

Distribusi-distribusi yang terdapat dalam keluarga eksponensial contohnya adalah distribusi *Bernoulli*, distribusi multinomial, dan distribusi normal, di mana setiap distribusi tersebut bekerja pada peubah yang berbeda-beda. Distribusi *Bernoulli* digunakan untuk peubah biner, distribusi multinomial digunakan untuk peubah nominal dan ordinal, dan distribusi normal digunakan untuk peubah bertipe kontinu.

### 2.3.1 Distribusi *Bernoulli*

Distribusi *Bernoulli* didefinisikan pada peubah acak biner (0 atau 1), distribusi ini menggunakan parameter  $\pi$ , di mana  $\pi = P(x = 1)$ . Distribusi

*Bernoulli* dapat ditulis sebagai:

$$f(x; \pi) = \pi^x(1 - \pi)^{1-x} \quad (2.3)$$

**Teorema 2.3.1.** Distribusi *Bernoulli* merupakan salah satu anggota dari keluarga eksponensial.

*Bukti.* Akan dibuktikan distribusi *Bernoulli* adalah distribusi dari keluarga eksponensial, maka distribusi *Bernoulli* dapat diubah menjadi persamaan (2.1), fungsi distribusi tersebut dapat dituliskan menjadi:

$$\begin{aligned} f(x; \pi) &= \pi^x \frac{1 - \pi}{(1 - \pi)^x} \\ &= (1 - \pi) \frac{\pi^x}{(1 - \pi)^x} \\ &= (1 - \pi) \exp \left\{ \ln \left\{ \frac{\pi}{1 - \pi} \right\}^x \right\} \\ &= (1 - \pi) \exp x \left\{ \ln \frac{\pi}{1 - \pi} \right\} \end{aligned}$$

maka, dari persamaan di atas, dapat ditentukan:

1.  $C(\theta) = 1 - \pi$ ,
2.  $t(x) = x$ ,
3.  $Q(\theta) = \ln \frac{\pi}{1 - \pi}$ ,
4. dan  $h(x) = 1$ .

□

### 2.3.2 Distribusi Multinomial

Distribusi multinomial didefinisikan pada peubah acak kategorik dengan pilihan kategori lebih dari dua, distribusi ini menggunakan parameter  $\pi$  untuk setiap kategori ke- $i$ , dengan  $i = 1, 2, \dots, K$ ,  $\pi$  akan bernilai satu jika objek memberikan respon untuk salah satu kategori, dan bernilai 0 untuk yang lain. Distribusi multinomial dapat ditulis sebagai:

$$f(x; \pi) = \prod_{i=1}^K \pi_i^{x_i} = \exp \left[ \sum_{i=1}^K x_i \ln \pi_i \right] \quad (2.4)$$

di mana  $x = (x_1, x_2, \dots, x_i)^T$  dan memenuhi  $\sum_{i=1}^K x_i = 1$ .

**Teorema 2.3.2.** Distribusi multinomial merupakan salah satu anggota dari keluarga eksponensial.

*Bukti.* Akan dibuktikan distribusi multinomial adalah distribusi dari keluarga eksponensial, maka distribusi multinomial dapat diubah menjadi persamaan (2.1), fungsi distribusi tersebut dapat dituliskan menjadi:

$$\begin{aligned} f(x; \pi) &= \exp \left[ \sum_{i=1}^K x_i \ln \pi_i \right] \\ &= \exp \left[ \sum_{i=1}^K x_i \ln \pi_i + \left( 1 - \sum_{i=1}^{K-1} x_i \right) \ln \left( 1 - \sum_{i=1}^{K-1} \pi_i \right) \right] \\ &= \left( 1 - \sum_{i=1}^{K-1} \pi_i \right) \exp \left[ \sum_{i=1}^{K-1} x_i \ln \left( \frac{\pi_i}{1 - \sum_{j=1}^{K-1} \pi_j} \right) \right] \end{aligned}$$

maka, dari persamaan di atas, dapat ditentukan:

1.  $C(\theta) = 1 - \sum_{i=1}^{K-1} \pi_i$ ,
2.  $t(x) = \sum_{i=1}^{K-1} x_i$ ,

$$3. Q(\theta) = \ln \left( \frac{\pi_i}{1 - \sum_{j=1}^{K-1} \pi_j} \right),$$

4. dan  $h(x) = 1$ .

□

### 2.3.3 Distribusi Normal

Distribusi normal atau sering disebut distribusi *Gauss* adalah distribusi peluang yang paling banyak digunakan dalam analisis statistika. Distribusi normal digunakan untuk peubah acak kontinu atau bersifat numerik.

**Definisi 2.3.2.** Ambil  $x$  adalah sebuah peubah acak kontinu, di mana nilai  $x$  adalah bilangan real  $R_x = \mathbb{R}$ , lalu ambil  $\mu \in \mathbb{R}$  dan  $\sigma \in \mathbb{R}$ . Dapat dikatakan bahwa  $x$  berdistribusi normal dengan rata-rata  $\mu$  dan *varians*  $\sigma^2$ , dengan fungsi kepadatan peluang:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} (x - \mu)^2 \right] \quad (2.5)$$

**Teorema 2.3.3.** Distribusi normal adalah sebuah distribusi dari keluarga eksponensial.

*Bukti.* Akan dibuktikan distribusi normal adalah distribusi dari keluarga eksponensial, maka distribusi normal dapat diubah menjadi persamaan (2.1), fungsi distribusi tersebut dapat dituliskan:

$$\begin{aligned} f(x; \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} (x - \mu)^2 \right] \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[ -\ln \sigma - \frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2} - \frac{\mu^2}{2\sigma^2} \right] \\ &= \frac{1}{\sigma \exp(\mu^2/2\sigma^2)} \frac{1}{\sqrt{2\pi}} \exp \left[ Q(\theta)t(x) - \ln \sigma - \frac{\mu^2}{2\sigma^2} \right]. \end{aligned}$$

maka, dari persamaan di atas, dapat ditentukan:

1.  $C(\theta) = \frac{1}{\sigma \exp(\mu^2/2\sigma^2)}$
2.  $Q(\theta) = [\mu/\sigma^2, -1/(2\sigma^2)]$
3.  $t(x) = [x, x^2]$ ,
4. dan  $h(x) = \frac{1}{\sqrt{2\pi}}$ .

□

## 2.4 Teorema Bayes

Suatu himpunan  $B_i$ , dengan  $i = 1, 2, \dots, m$ ,  $B_i$  dikatakan partisi dari ruang sampel  $S$  jika:

1.  $B_j \cap B_k = \emptyset$ , untuk  $j \neq k$  di mana  $j = 1, 2, \dots, m$  dan  $k = 1, 2, \dots, m$ ,
2.  $\cup_{i=1}^m B_i = S$ ,
3.  $P(B_i) > 0$  untuk  $\forall_i$ .

Bila  $A$  merupakan peristiwa bagian dari  $S$  sehingga  $A = \cup_{i=1}^m (A \cap B_i)$  di mana masing-masing  $(A \cap B_i)$  adalah saling lepas, maka

$$P(A) = P(\cup_{i=1}^m (A \cap B_i)) = \sum_{i=1}^m P(A \cap B_i)$$

dan dengan menggunakan peluang bersama diperoleh:

$$P(A) = \sum_{i=1}^m P(A|B_i)P(B_i) \tag{2.6}$$

secara umum, untuk semua  $i$  berlaku:

$$P(A \cap B_i) = \sum_{i=1}^m P(A|B_i)P(B_i)$$

pembagian dengan  $P(A)$  menghasilkan:

$$\frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A)} \quad (2.7)$$

dengan mensubstitusikan persamaan-persamaan di atas, didapatkan:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^m P(A|B_i)P(B_i)} \quad (2.8)$$

Persamaan di atas merupakan formula dari Teorema Bayes yang menjelaskan peluang terjadinya kejadian ke  $B_i$  setelah terjadi kejadian  $A$ .

## 2.5 Uji Normalitas

Uji normalitas berguna untuk membuktikan data dari sampel yang dimiliki berasal dari populasi berdistribusi normal. Banyak cara yang dapat dilakukan untuk membuktikan suatu data berdistribusi normal atau tidak, diantaranya dengan uji normalitas *Shapiro-Wilk*.

Uji normalitas *Shapiro-Wilk* digunakan untuk jumlah data yang lebih dari 30 ( $N > 30$ ). Menurut Razali dan Y.B. Wah (2011) uji *Shapiro-Wilk* merupakan uji normalitas yang paling kuat dibandingkan uji *Kolmogorov-Smirnov*, *Anderson-Darling*, dan *Liliefors* karena sangat sensitif terhadap data yang tidak normal, serta memberikan *p-value* yang lebih besar dengan berbagai tingkat signifikansi.

*Shapiro-Wilk* menggunakan data asli yang belum diolah dalam tabel distribusi frekuensi. Data diurut dari yang terkecil hingga terbesar. Tahapan hipotesis untuk uji ini adalah:

1. Menentukan hipotesis,

$H_0$  : data berasal dari populasi yang berdistribusi normal,

$H_1$  : data tidak berasal dari populasi yang berdistribusi normal.

2. Menentukan  $\alpha$ ,

3. Statistik uji,

$$W = \frac{1}{D} \left[ \sum_{i=1}^N a_i (X_{N-i+1} - X_i) \right]^2 \quad (2.9)$$

dengan;

$$D = \sum_{i=1}^N (X_i - \bar{X})^2 \quad (2.10)$$

Normalisasi untuk statistik  $W$  yang diperkenalkan oleh Royston (1992) yaitu:

$$Z_W = \frac{\ln(1 - W) - \mu}{\sigma} \quad (2.11)$$

untuk  $12 \leq N \leq 2000$ , dengan:

$$\mu = -1,5861 - 0,31082 \ln(N) - 0,083751(\ln(N))^2 + 0,0038915(\ln(N))^3 \quad (2.12)$$

dan;

$$\sigma = \exp(-0,4803 - 0,082676 \ln(N) - 0,0030302(\ln(N))^2), \quad (2.13)$$

4. Kriteria pengujian,

$H_0$  diterima jika nilai  $Z_W < Z_\alpha$ ,

$H_1$  ditolak jika nilai  $Z_W > Z_\alpha$ ,

dengan  $Z_\alpha$  diperoleh dari tabel distribusi normal standar,  $X_i$  merupakan data ke- $i$ ,  $\bar{X}$  rata-rata dari data, dan  $a_i$  koefisien dari test *Shapiro-Wilk*.

Untuk normalisasi data yang tidak normal, perlu dilakukan sebuah transformasi. Menurut Shinta (2014), salah satu transformasi yang dapat dilakukan adalah dengan cara *me-lognormal* kan keseluruhan data atau mengubah data dengan sebuah persamaan tertentu, hingga asumsi normalitas terpenuhi (*p-value Shapiro-Wilk*  $> 0,05$ ).

## 2.6 *Latent Class Cluster Analysis*

Model *LCCA* diperkenalkan pada tahun 1963 oleh Lazarsfeld sebagai suatu alat untuk mengklasifikasikan berdasarkan pada peubah dikotomis yang diamati menggunakan model distribusi kepadatan peluang dan mencari penaksiran parameter fungsi kepadatan peluang tersebut. Pada tahun 1968, Lazarsfeld dan Henry memperkenalkan model *LCCA* sebagai suatu cara untuk memformulasikan peubah laten dari survey yang diamati dengan peubah dikotomis yaitu ada tidaknya ciri-ciri tertentu pada suatu objek. Model *LCCA* mengasumsikan bahwa peubah laten adalah peubah kategorik (Vermunt dan Magidson, 2004). Pada tahun 2004, Moustaki dan Papageorgiou mengembangkan model *LCCA* yang tidak hanya untuk peubah kategorik, melainkan gabungan dengan peubah kontinu.

Kegunaan *Latent Class Analysis* ini antara lain:

1. Pembentukan gerombol (*clustering*) sekelompok data berindikator campuran dengan menggunakan *Latent Class Cluster Analysis*,
2. Pemilihan peubah dan pembentukan skala menggunakan *Latent Class Factor and Choices Models*,
3. Penaksiran peubah tak bebas menggunakan *Latent Class Regression Analysis*.

Model *LCCA* digunakan untuk mengklasifikasikan individu atau objek ke dalam kelompok/kelas yang berbeda berdasarkan pada respon untuk sekumpulan peubah yang diamati. Tujuan dari *LCCA* adalah pertama untuk menentukan jumlah kelas laten optimal yang terbentuk, lalu mengalokasikan responden / objek ke dalam kelas laten (Moustaki dan Papageorgiou, 2004).

*LCCA* sebagai suatu metode penggerombolan yang menggunakan model statistik mempunyai beberapa kelebihan sebagai berikut (Vermunt dan Magidson, 2002):

1. Kriteria pemilihan gerombol lebih jelas dan pendekatannya memasukkan pengujian statistik yang kuat.
2. *LCCA* sangat fleksibel untuk bentuk distribusi yang sederhana atau yang rumit, dapat digunakan untuk berbagai jenis peubah teramati di dalam gerombol.
3. Tidak ada keputusan yang harus dibuat tentang penskalaan peubah teramati.

4. Ada kriteria yang lebih formal untuk membuat keputusan tentang jumlah gerombol.

Hidayat (2009) membandingkan hasil klasifikasi menggunakan *LCCA* dan *K-Means* menggunakan data simulasi. Hasilnya, *LCCA* mempunyai beberapa kelebihan dibandingkan analisis gerombol *K-Means* yaitu pada *LCCA* ada patokan yang jelas berapa banyak kelompok yang akan dibentuk (menggunakan nilai BIC, dan AIC), dapat diaplikasikan pada data dengan berbagai skala pengukuran, tidak memerlukan standarisasi peubah, tidak memerlukan analisis lanjutan karena karakteristik objek dalam satu kelompok dapat dilihat dari penaksiran parameternya secara langsung.

Dalam Vermunt dan Magidson (2004), beberapa peneliti telah menempatkan model *LCCA* ke dalam kerangka *log-linier*, salah satunya Vermunt dan Magidson (2002) yang mengharuskan semua peubah adalah berasal dari peubah kategorik. *LCCA* dalam bentuk yang lebih klasik telah dibahas untuk penggerombolan peubah biner, nominal atau metrik di Bartholomew dan Knott (1999) dan untuk penggerombolan dengan data campuran di Everitt (1988), Everitt dan Merette (1990) dalam Moustaki (2004) dan Moustaki (1996).

Dalam penaksiran parameter model *LCCA* untuk menganalisis hubungan antara sekelompok peubah teramati campuran menggunakan satu atau lebih peubah laten, Moustaki (1996) menggunakan algoritma EM yang secara komputasi sangat efisien. Kelompok peubah teramati berisi peubah metrik (kontinu atau diskrit) dan biner.

Berikut ini adalah tabel perbedaan analisis gerombol menggunakan *K-Means* dan menggunakan *LCCA*:

Tabel 2.1: Perbedaan *K-Means* dengan *LCCA*

Perbedaan	<i>K-Means</i>	<i>LCCA</i>
Jumlah gerombol	Ditentukan terlebih dahulu	Dicari banyaknya gerombol yang paling optimal
Ukuran yang digunakan	Jarak ( <i>distance</i> ) biasanya dengan jarak <i>euclid</i>	Distribusi peluang bersama
Jenis data	Hanya menggunakan data kontinu saja	Dapat digunakan berbagai jenis skala data

Penelitian yang dilakukan Moustaki dan Papageorgiou (2004) memperluas *LCCA* untuk peubah campuran biner dan metrik dengan mengakomodasi semua jenis data (termasuk ordinal dan nominal) dan menggunakannya dalam *archaeometry* untuk mengklasifikasikan penemuan/objek arkeologi menjadi beberapa kelompok. Penaksiran untuk model digunakan *maximum likelihood* dengan algoritma EM.

### 2.6.1 Model *Latent Class Cluster Analysis* untuk Peubah Kategorik

Pada prinsipnya, *LCCA* tidak jauh berbeda dengan analisis gerombol *K-Means*, yaitu mengelompokkan sejumlah objek yang terdekat dengan pusat kelompoknya sehingga jarak setiap individu ke pusat kelompok dalam satu kelompok adalah minimum. Perbedaannya adalah, jika pada *K-means* digunakan pendekatan jarak untuk menyatakan kedekatan objek pada pusatnya sedangkan pada *LCCA*, untuk menyatakan bahwa satu objek dekat dengan pusatnya didasarkan pada peluang pengelompokan yang ditaksir melalui metode *maximum likelihood* (Vermunt dan Magidson, 2002). Jika dalam analisis gerombol *K-Means* jarak seti-

ap objek yang terdekat dengan pusat gerombol dikumpulkan ke dalam satu gerombol, maka dalam *LCCA* menggunakan peluang terdekat dengan pusat gerombol yang akan dibuat.

Pada model *LCCA* akan dibentuk sebanyak  $J$  gerombol terhadap  $N$  objek yang terdiri dari  $P$  peubah campuran, maka proses penggerombolan menggunakan distribusi peluang bersama dari masing-masing peubah acak. Untuk setiap gerombol  $j$  dihubungkan dengan peluang  $\eta_j$  (Moustaki dan Papageorgiu, 2004). Sebaran untuk masing-masing peubah acak dalam setiap gerombol adalah:

$$f(x_h) = \prod_{j=1}^J \eta_j g(x_{ih}|j) \quad (2.14)$$

dimana:

1.  $f(x_h)$  adalah fungsi peluang bersama untuk peubah-peubah pengamatan pada objek ke- $h$ ,
2.  $g(x_{ih}|j)$  adalah fungsi distribusi peluang peubah acak ke  $i$  di dalam gerombol ke  $j$ , untuk objek ke  $h$ ,
3.  $\eta_j$  adalah peluang awal kelas laten atau gerombol ke  $j$ , dimana  $\eta_j = \frac{n_j}{N}$ ,
4.  $J$  adalah banyaknya gerombol ( $j = 1, 2, \dots, J$ ), dimana  $\sum_{j=1}^J \eta_j = 1$ .

Berikut ini adalah masing-masing model *LCCA* untuk masing-masing peubah kategorik:

## 1. Peubah Biner

Untuk peubah biner yang mengambil nilai-nilai 0 dan 1 maka mengikuti distribusi *Bernoulli* yaitu sebagai berikut:

$$g(x_{ih}|j) = \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \quad (2.15)$$

dimana:

- (a)  $\pi_{ij}$  adalah peluang objek akan berada di kelas ke- $j$  untuk peubah ke- $i$  ( $\pi_{ij} = \frac{n_i}{n_j}$ ),
- (b)  $B$  adalah banyaknya peubah biner yang digunakan dalam penelitian.

maka model *LCCA* untuk peubah biner adalah:

$$f(x_h) = \sum_{j=1}^J \eta_j g(x_{ih}|j) = \sum_{j=1}^J \eta_j \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \quad (2.16)$$

## 2. Peubah Nominal

Untuk peubah nominal, dalam kasus *polytomous* peubah indikator  $x_i$  digantikan oleh sebuah nilai vektor fungsi indikator yang didefinisikan  $x_{i(s)} = 1$  jika respon jatuh dalam kategori  $s$ , untuk  $s = 1, 2, \dots, S$  dan  $x_{i(s)} = 0$  untuk lainnya, di mana  $S$  menunjukkan jumlah kategori peubah ke- $i$  dan  $\sum_{s=1}^S x_{i(s)} = 1$ .

Pola respon dari objek ke- $h$  ditulis sebagai  $x'_h = (x_{1h}, x_{2h}, \dots, x_{ph})$  sebagai dimensi dari  $S$ . Peluang respon tunggal  $\pi_{ij}$  dari kasus biner diganti dengan satu himpunan fungsi  $\pi_{ij(s)} (s = 1, 2, \dots, S)$  di mana  $\sum_{s=1}^S \pi_{ij(s)} = 1$ . Distribusi untuk peubah nominal ini menggunakan distribusi multinomial,

yaitu:

$$g(x_{ih}|j) = \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \quad (2.17)$$

dimana:

- (a)  $\pi_{ij(s)}$  adalah peluang objek akan berada di kelas  $j$  untuk kategori ke- $s$  dan peubah ke- $i$  ( $\pi_{ij(s)} = \frac{n_{i(s)}}{n_j}$ ).
- (b)  $s$  adalah kategori jawaban yang digunakan pada peubah ke- $i$  ( $s = 1, 2, \dots, S$ ).
- (c)  $J$  adalah banyaknya gerombol yang terbentuk ( $j = 1, 2, \dots, J$ ).
- (d)  $L$  adalah banyaknya peubah nominal yang digunakan

Model *LCCA* untuk peubah nominal memiliki bentuk:

$$f(x_h) = \sum_{j=1}^J \eta_j g(x_{ih}|j) = \sum_{j=1}^J \eta_j \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \quad (2.18)$$

### 3. Peubah Ordinal

Untuk peubah ordinal, peubah bebas  $x_{ih}$  didefinisikan seperti pada peubah nominal. Kategori (urutan)  $R$  memiliki peluang  $\pi_{ij(1)}$ ,  $\pi_{ij(2)}$ , ...,  $\pi_{ij(R)}$ , untuk menjaga urutan dari peubah  $x_i$ , peluang kumulatif dari respon dimodelkan seperti di bawah ini dan memasukkan kategori  $s$  sebagai:

$$\gamma_{ij(s)} = \pi_{ij(1)} + \pi_{ij(2)} + \dots + \pi_{ij(s)} \quad (2.19)$$

di mana  $j = 1, \dots, J$  dan  $s = 1, \dots, R$ .

Untuk peubah teramati  $x_{ih}$ , fungsi distribusi mengikuti distribusi multinomial yakni sebagai berikut:

$$g(x_{ih}|j) = \prod_{i=1}^O \prod_{s=1}^R \pi_{ij(s)}^{x_{ih(s)}} = \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \quad (2.20)$$

dimana:

- (a)  $\pi_{ij(s)}$  adalah peluang objek akan berada di kelas  $j$  untuk kategori ke- $s$  pada peubah ke- $i$  ( $\pi_{ij(s)} = \frac{n_{i(s)}}{n_j}$ ),
- (b)  $s$  adalah kategori jawaban yang digunakan pada peubah ke- $i$  ( $s = 1, 2, \dots, R$ ),
- (c)  $J$  adalah banyaknya kelas atau gerombol yang terbentuk ( $j = 1, 2, \dots, J$ ),
- (d)  $\gamma_{ij(s)}$  adalah peluang kumulatif respon akan berada di kelas  $j$  untuk kategori ke- $s$  pada peubah ke- $i$ .
- (e)  $O$  adalah banyaknya peubah ordinal yang digunakan dalam penelitian.

Model *LCCA* untuk peubah ordinal memiliki bentuk:

$$f(x_h) = \sum_{j=1}^J \eta_j g(x_{ih}|j) = \sum_{j=1}^J \eta_j \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \quad (2.21)$$

### 2.6.2 Model *Latent Class Cluster* untuk Peubah Kontinu

Menurut Moustaki dan Papageorgiou (2004), *LCCA* sebagai suatu alat untuk analisis gerombol dengan indikator kontinu, diasumsikan peubah-peubah kontinu tersebut berdistribusi normal. Bentuk model *LCCA* untuk peubah kon-

tinu adalah sebagai berikut:

$$f(x_h) = \sum_{j=1}^J \eta_j f(x_{ih}; \mu_j, \sigma_i) \quad (2.22)$$

di mana  $f(x_h)$  menggambarkan distribusi peluang dari objek ke  $x_h$ ,  $f(x_{ih}; \mu_j, \sigma_i)$  adalah distribusi dari  $x_{ih}$  didalam kelas  $j$  untuk  $\mu_j$  dan  $\sigma_i$ , dengan  $\mu_j$  melambangkan vektor rata-rata kelas ke- $j$ , dan  $\sigma_i$  adalah *varians* peubah ke- $i$ .

Pada kasus di mana peubah bebas  $x_{ih}$  berdistribusi normal, sebaran  $g(x_{ih}|j)$  adalah:

$$g(x_{ih}|j) = \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \quad (2.23)$$

di mana:

1.  $\mu_{ij}$  adalah rata-rata peubah ke- $i$  dalam kelas ke- $j$ ,
2.  $\sigma_i^2$  adalah ragam peubah ke- $i$  yang dianggap konstan pada setiap kelas.

Model *LCCA* untuk peubah kontinu memiliki bentuk:

$$\begin{aligned} f(x_h) &= \sum_{j=1}^J \eta_j g(x_{ih}|j) \\ &= \sum_{j=1}^J \eta_j \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \end{aligned} \quad (2.24)$$

## 2.7 Penaksiran Maksimum *Likelihood*

Misalkan  $\Theta = \{\theta_1, \theta_2, \dots, \theta_i\}$  merupakan suatu himpunan parameter fungsi kepadatan peluang bersama dari beberapa distribusi, maka untuk mencari nilai taksiran untuk  $\Theta$  salah satunya dengan metode maksimum *likelihood*.

Untuk melakukan penaksiran dengan metode ini, ada beberapa tahapan yang harus dilakukan, yakni sebagai berikut:

1. Penentuan fungsi kepadatan peluang bersama.

Fungsi kepadatan peluang bersama dari  $x_1, x_2, \dots, x_{ih}$ , dengan setiap peubah acak saling bebas  $x_i$  memiliki parameter  $\theta_i$ , dapat dinyatakan sebagai:

$$f(x_i; \theta_1, \theta_2, \dots, \theta_i) = f(x_i; \theta_1) \cdot f(x_i; \theta_2) \cdot \dots \cdot f(x_i; \theta_i) \quad (2.25)$$

2. Penentuan fungsi *likelihood*.

Fungsi *likelihood* adalah fungsi kepadatan peluang bersama dari  $x_1, x_2, \dots, x_i$  yang dianggap sebagai fungsi dari  $\theta$ , sehingga persamaan 2.25 dapat ditulis sebagai berikut:

$$\begin{aligned} L(\theta_1, \theta_2, \dots, \theta_n) &= f(x_i; \theta_1, \theta_2, \dots, \theta_i) \\ &= f(x_i; \theta_1) \cdot f(x_i; \theta_2) \cdot \dots \cdot f(x_i; \theta_i) \\ &= \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_i) \end{aligned}$$

3. Penaksiran nilai untuk  $\theta$ .

Dalam metode penaksiran maksimum *likelihood*, taksiran dari  $\theta$ , sebut saja  $\hat{\theta}$ , diperoleh dengan cara memaksimumkan fungsi *likelihood*  $L(\theta)$ . Nilai  $\theta$  yang memaksimumkan  $L(\theta)$  sama dengan nilai  $\theta$  yang memaksimumkan  $\ln L(\theta)$ , sebut saja  $l(\theta)$ , sehingga persamaan  $l(\theta)$  menjadi:

$$l(\theta) = \ln L(\theta) = \ln \prod_{i=1}^n f(x_i; \theta_i)$$

sehingga didapat:

$$l(\theta) = \sum_{i=1}^n \ln f(x_i; \theta_i) \quad (2.26)$$

Nilai  $\theta$  yang memaksimumkan  $l(\theta)$  dapat diperoleh dengan mencari solusi dari persamaan

$$\frac{\partial}{\partial \theta_i} l(\theta_i) = 0, \quad i = 1, \dots, m$$

Namun seringkali pada persamaan tersebut tidak dapat diselesaikan dengan cara analitik, melainkan dapat menggunakan prosedur numerik.

## 2.8 Algoritma EM (*Expectation-Maximization*)

Algoritma EM adalah salah satu metode iteratif yang biasanya digunakan untuk mendapatkan solusi dari persamaan *likelihood*. Dalam algoritma ini terdapat istilah data teramati (*observed*) dan data tidak teramati (*latent*). Menurut Dempster *et al* (1997) algoritma EM terdiri dari dua tahap, yaitu tahap E (Ekspektasi) dan tahap M (Maksimisasi).

Misalkan  $Y$  adalah suatu peubah laten dengan  $J$  kategori dan  $X_1, X_2, \dots, X_n$  adalah peubah teramati yang mempunyai fungsi kepadatan peluang bersama  $f(x; \theta)$ . Fungsi *log-likelihood* dari  $X$  dapat dinyatakan sebagai berikut:

$$l(x, \theta) = \ln[f(x; \theta)] \quad (2.27)$$

Misalkan  $f(x, y, \theta)$  adalah fungsi kepadatan peluang bersama dari  $X$  dan  $Y$ , dengan  $\theta$  adalah parameter dalam model, maka salah satu cara untuk mencari taksiran  $\theta$  yang memaksimumkan fungsi *likelihood* dari  $X$  adalah dengan meng-

gunakan algoritma EM. Prinsip dari algoritma EM ini dapat dibagi menjadi dua bagian, sebagai berikut:

1. *E-Step*

*E-step* dilakukan untuk mencari  $E[\ln[f(y, x, \theta_t)]|x, \hat{\theta}_{t-1}]$ , di mana:

$\hat{\theta}_{t-1}$  adalah taksiran  $\theta$  pada iterasi ke- $t - 1$ ,  $t = 1, 2, \dots$

$\theta_t$  adalah nilai  $\theta$  pada iterasi ke- $t$ .

$\theta_0$  adalah suatu nilai taksiran awal yang diberikan.

2. *M-Step*

Pada *M-step*, maksimasi  $E[\ln[f(y, x, \theta_t)]|x, \hat{\theta}_{t-1}]$  dengan cara menurunkannya terhadap  $\theta_t$ , untuk mendapatkan taksiran  $\theta_t$  pada iterasi ke- $t$ , yang disebut  $\hat{\theta}_t$ .

Proses *E-step* dan *M-step* ini akan dilakukan terus secara iteratif sampai didapatkan suatu estimasi untuk  $\theta$  yang konvergen atau  $|\hat{\theta}_t - \hat{\theta}_{t-1}|$  cukup kecil atau mendekati nol. Iterasi algoritma EM akan meningkatkan nilai  $l(x, \theta)$  pada setiap iterasinya (Welling, 2001).

Selain menggunakan pendugaan maksimum *likelihood*, estimasi parameter dapat pula digunakan dengan metode konvergensi komputasi yang terdapat dalam *software* LatentGold 4.0. Kriteria konvergensi menggunakan bentuk sebagai berikut:

$$CC = \sum_{p=1}^P \left| \frac{\hat{\theta}_p^t - \hat{\theta}_p^{t-1}}{\hat{\theta}_p^{t-1}} \right| \quad (2.28)$$

dengan  $P$  menunjukkan banyaknya parameter yang akan diestimasi.

Iterasi ke- $t$  akan berhenti jika pergantian dalam *log-posterior* dalam *output* LatentGold 4.0 memiliki nilai lebih kecil dari  $10^{-12}$  (Vermunt dan Magidson, 2004).

## 2.9 Pemilihan Model Terbaik

Untuk pemilihan model terbaik dalam *LCCA* dapat dilakukan dengan uji statistik, diantaranya AIC (*Akaike's Information Criterion*) dan BIC (*Bayesian Information Criterion*).

### 2.9.1 AIC (*Akaike's Information Criterion*)

Menurut Fraley dan Raftery (1998) dalam Nainggolan (2008), kriteria dalam pemilihan model terbaik dalam *LCCA* yaitu dengan menggunakan kriteria informasi Akaike (AIC, *Akaike's Information Criterion*) dengan rumus:

$$AIC = -2l(\hat{\theta}) + 2q \quad (2.29)$$

dengan  $l(\hat{\theta})$  adalah nilai maksimum *log-likelihood* dari model dan  $q$  adalah banyaknya parameter dalam model.

Secara umum, semakin kecil nilai AIC maka model yang dipakai semakin cocok. Model yang dianggap terbaik adalah model yang memiliki nilai AIC minimum. Namun demikian, dengan pertimbangan aspek lain, perbedaan AIC yang tidak terlalu besar mungkin dapat diabaikan (Tirta, 2009).

### 2.9.2 BIC (*Bayesian Information Criterion*)

Selain menggunakan kriteria statistik AIC (*Akaike's Information Criterion*) untuk memilih model yang terbaik bisa digunakan juga BIC (*Bayesian Information Criterion*). Menurut Nainggolan (2008), *Bayesian Information Criterion* (BIC) adalah sebuah statistik untuk memilih model, nilai BIC mencerminkan peningkatan nilai jumlah kuadrat residual dan jumlah parameter dari

model yang digunakan. Variasi yang tidak bisa dijelaskan oleh model mengakibatkan peningkatan BIC. Model dengan nilai BIC lebih kecil menunjukkan bahwa model yang dihasilkan lebih bisa menjelaskan variasi dari data. BIC yang dihitung berdasarkan *log-likelihood* dirumuskan sebagai berikut:

$$BIC = -2l(\hat{\theta}) + \ln(N)q \quad (2.30)$$

dengan  $l(\hat{\theta})$  adalah nilai maksimum fungsi *log-likelihood* yang ditaksir,  $N$  adalah banyaknya observasi dan  $q$  adalah banyaknya parameter.

## 2.10 Uji Kelayakan Model

Model yang dipilih perlu diketahui seberapa besar tingkat kelayakan dengan menggunakan nilai statistik *Wald*. Menurut Vermunt dan Magidson (2004), signifikansi dari sekelompok parameter dapat diuji dengan menggunakan statistik *Wald*. Tahapan uji hipotesis untuk uji ini adalah:

1. Menentukan hipotesis,

$$H_0 : \theta = \theta_0 \text{ (model signifikan dari } \hat{\theta} \text{ yang tidak nol),}$$

$$H_1 : \theta \neq \theta_0 \text{ (model tidak signifikan)}$$

2. Menentukan  $\alpha$ ,

3. Statistik uji,

$$W^2 = (\hat{\theta} - \theta_0)' [cov(\hat{\theta})]^{-1} (\hat{\theta} - \theta_0) \quad (2.31)$$

4. Kriteria pengujian,

$$H_0 \text{ diterima jika nilai } p\text{-value} > \alpha,$$

$$H_0 \text{ ditolak jika nilai } p\text{-value} < \alpha.$$

## 2.11 Uji Asumsi Kebebasan Lokal

Model *LCCA* memiliki asumsi bahwa setiap peubah kategorik yang berada dalam satu gerombol haruslah saling bebas. Asumsi ini harus terpenuhi untuk mendapatkan hasil klasifikasi yang akurat. Untuk menguji asumsi ini digunakan tabel kontingensi *chi-square* untuk melihat hubungan pasangan peubah kategorik tersebut.

Pelanggaran asumsi ini dapat diperiksa menggunakan statistik uji *Bivariate Residual (BVR)*. Hipotesis untuk kasus ini adalah sebagai berikut:

$H_0$  : Kebebasan lokal pada model *LCCA* terpenuhi.

$H_1$  : Kebebasan lokal pada model *LCCA* tidak terpenuhi.

Kesimpulan  $H_0$  ditolak pada saat nilai *BVR* cukup besar, yaitu lebih besar dari  $\chi^2_{(\alpha; db=1)}$  (Vermunt dan Magidson, 2004). Nilai *BVR* dapat diperoleh dengan menggunakan rumus:

$$BVR = \frac{\chi^2}{db} = \frac{\sum_{i=1}^b \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}}{db} \quad (2.32)$$

dimaan  $O_{ij}$  merupakan frekuensi observasi,  $E_{ij} = \frac{n_i m_j}{N}$  merupakan frekuensi harapan dan  $db = (b - 1)(k - 1)$  adalah derajat bebas dengan  $b$  adalah jumlah baris dan  $k$  adalah jumlah kolom yang merupakan kategori dalam peubah.

Menurut Vermunt dan Magidson (2004), terdapat alternatif solusi apabila asumsi kebebasan lokal tidak terpenuhi, pertama dengan menghapus salah satu peubah, namun cara ini tidak digunakan ketika banyak peubah yang tidak saling bebas. Kedua dapat meningkatkan jumlah kategori dalam suatu peubah, misalkan dalam satu peubah terdapat dua kategori yaitu tinggi dan rendah, menjadi tinggi, sedang, dan rendah. Ketiga, menurut Vermunt dan Magidson (2004)

dengan penambahan satu atau lebih pengaruh langsung (*direct effect*) dengan menambahkan gabungan dua peubah yang tidak saling bebas pada model *LCCA* sampai didapatkan nilai BVR yang memenuhi asumsi.

## 2.12 Klasifikasi Objek

Klasifikasi objek ke dalam kelompok oleh model didasarkan pada nilai taksiran peluang pada objek ke- $h$  atau  $\hat{h}(j|x_h)$ . Kemungkinan sebuah objek menjadi anggota kelompok ke- $j$ , dimana  $j$  adalah kelompok dan  $x_h$  adalah objek ke- $h$  adalah sebagai berikut:

$$\hat{h}(j|x_h) = \max\{\hat{h}(J = j|x_h)\}, \quad j = 1, 2, \dots, J \quad (2.33)$$

atau dengan kata lain suatu objek masuk ke dalam kelompok ke- $k$  dengan peluang pengelompokan tertinggi. Dengan nilai dari  $\hat{h}(j|x_h)$  dapat diperoleh informasi tentang ke dalam kelompok manakah sebuah objek dapat diklasifikasikan.

## 2.13 Keragaman Gerombol

Pada dasarnya dalam setiap penggerombolan terdapat dua keragaman, yaitu keragaman dalam gerombol (*Sum of Square Within Cluster / SSW*), dan keragaman antar gerombol (*Sum of Square Between Cluster / SSB*). *Sum of square within cluster* (SSW) atau keragaman dalam gerombol adalah suatu nilai keragaman pada setiap elemen pada suatu gerombol yang merupakan nilai *varians* dari gerombol tersebut, semakin rendah nilai SSW mengartikan bahwa gerombol tersebut semakin baik atau memiliki homogenitas internal yang tinggi. *Sum of*

*square between cluster* (SSB) adalah nilai keragaman antar gerombol, semakin tinggi nilai SSB mengartikan bahwa gerombol tersebut semakin baik dan tingkat perbedaan antar gerombol semakin tinggi.

### 1. Keragaman untuk data kontinu

Bunkers *et al* (1996) dalam Dewi dkk. (2012) menerangkan persamaan rumus sebagai berikut untuk menghitung SSW dan SSB data dengan peubah kontinu, yakni sebagai berikut:

$$SSW = \frac{1}{J} \sum_{j=1}^J S_j \quad (2.34)$$

$$SSB = \left[ \frac{1}{J-1} \sum_{j=1}^J (\bar{x}_j - \bar{x})^2 \right] \quad (2.35)$$

dimana:

$SSW$  adalah keragaman dalam gerombol,

$SSB$  adalah keragaman antar gerombol,

$S_j$  adalah simpangan baku gerombol ke- $j$

$$S_j = \sqrt{\frac{\sum_{h=i}^N (x_i - \bar{x})^2}{N-1}} \quad (2.36)$$

$\bar{x}_j$  adalah rata-rata gerombol ke- $j$ ,

$\bar{x}$  adalah rata-rata total,

$J$  adalah banyaknya gerombol.

### 2. Keragaman untuk data kategorik

Light dan Margolin (1997) dalam Dewi dkk. (2012) menjelaskan rumus untuk menghitung SSW dan SSB pada peubah kategorik, yaitu se-

bagai berikut:

$$SSW = \left[ \frac{1}{N - J} \left( \frac{N}{2} - \frac{1}{2} \sum_{j=1}^J \frac{1}{n_j} \sum_{s=1}^S n_s^2 \right) \right]^{\frac{1}{2}} \quad (2.37)$$

$$SSB = \left[ \frac{1}{J - 1} \left( \frac{1}{2} \left( \sum_{j=1}^J \frac{1}{n_j} \sum_{s=1}^S n_{sj}^2 \right) - \frac{1}{2N} \sum_{s=1}^S n_s^2 \right) \right]^{\frac{1}{2}} \quad (2.38)$$

dimana:

$SSW$  adalah keragaman dalam gerombol,

$SSB$  adalah keragaman antar gerombol,

$n_s$  adalah banyaknya objek pada kategori ke- $s$ ,  $s = 1, 2, \dots, S$ ,

$n_j$  adalah banyaknya objek pada gerombol ke- $j$ ,  $j = 1, 2, \dots, J$ ,

$n_{sj}$  adalah banyaknya objek pada kategori ke- $s$ , gerombol ke- $j$ .

Hasil penggerombolan yang baik adalah jika objek dalam gerombol yang sama memiliki keragaman yang rendah sedangkan objek antar gerombol memiliki tingkat keragaman yang tinggi. Dengan kata lain, objek dalam satu gerombol memiliki tingkat kemiripan yang tinggi dan objek berbeda gerombol memiliki tingkat kemiripan yang rendah (Serban dan Grigoreta, 2006 dalam Dewi dkk, 2012).

Pengelompokan pada *LCCA* dilakukan dengan menggunakan *software* Latent GOLD v4.0. Penentuan banyaknya gerombol yang terbentuk dapat ditentukan dengan memilih nilai BIC terkecil, berdasarkan banyak nya gerombol yang terbentuk pada *LCCA*, pada setiap datanya kemudian dilakukan perhitungan keragaman, baik keragaman dalam gerombol ( $SSW$ ) maupun keragaman antar gerombol ( $SSB$ ). Nilai keragaman yang diperoleh tersebut kemudian dihitung rasionya dengan cara membagi nilai  $SSW$  dengan  $SSB$  untuk masing-masing

tipe peubah.

Pembagian nilai SSW dan SSB ini diharapkan dapat semakin memperkecil nilai rasio keragamannya, dikarenakan pada nilai SSW dibuat sekecil mungkin, dan nilai SSB dibuat sebesar mungkin pada gerombol optimal yang telah dipilih, sehingga akan diperoleh nilai rasio keragaman yang paling rendah atau kriteria gerombol maksimal dalam pengelompokan objek telah tercapai.

Nilai akurasi sebuah gerombol yang baik dalam menggerombolkan  $N$  objek menurut Umah (2014), adalah sebagai berikut:

$$Akurasi(\%) = \left( 1 - \frac{\sum_{i=1}^P \frac{SSW_i}{SSB_i}}{P} \right) \times 100\% \quad (2.39)$$

Kriteria nilai rasio keragaman yang baik dalam penggerombolan objek adalah kurang dari satu atau tingkat akurasinya mendekati 100%, semakin kecil nilai rasio keragaman atau mendekati nol, maka gerombol yang terbentuk memiliki homogenitas internal dan homogenitas eksternal yang tinggi (Umah, 2014).

## 2.14 Karakteristik *Smartphone*

*Smartphone* merupakan salah satu media teknologi yang paling cepat mengalami perkembangan. Dalam beberapa tahun belakangan, produk-produk *smartphone* mengalami inovasi yang luar biasa, terlebih dalam aspek *hardware*. Inovasi inilah yang kemudian menarik para konsumen untuk menjadikan *smartphone* seakan kebutuhan utama mereka.

Semakin tingginya tingkat antusiasme masyarakat, menjadikan produsen *smartphone* berlomba-lomba memasarkan produk mereka, sehingga banyak produk-produk di masyarakat dengan karakteristik yang beragam. Hal ini kemudian men-

jadikan pangsa pasar *gadget* sebagai salah satu lahan bisnis yang diminati dan sangat berkembang.

Beragamnya produk-produk *smartphone* di kalangan masyarakat, membuat konsumen terkadang sulit dalam menentukan pilihan yang akan mereka beli. Ditambah lagi, belum adanya klasifikasi yang jelas mengenai produk-produk yang ada di masyarakat, terkait kelompok-kelompok produk yang bagaimanakah yang terdapat di masyarakat. Dengan adanya pengklasifikasian tentunya dapat memudahkan konsumen dalam menentukan pilihan, serta memberikan informasi kepada produsen tentang produk-produk yang seperti apa yang sekelas dengan produk buatan mereka sehingga dapat menambah daya saing antar produsen.

Menurut Elan (2015) terdapat beberapa hal yang menjadikan sebuah produk diminati masyarakat, yang seringkali dijadikan patokan dalam melakukan keputusan pembelian suatu produk yakni sebagai berikut:

1. OS dan *processor* yang digunakan

Sistem operasi (OS) adalah seperangkat program yang mengelola sumber daya perangkat keras (*hardware*) dan menyediakan layanan umum untuk aplikasi perangkat lunak. Terdapat empat jenis sistem operasi *smartphone* yang berkembang di pasaran akhir-akhir ini yaitu Android, iOS, BlackBerry, dan Windows Phone. Setiap sistem operasi memiliki kinerja dan *user interface* tersendiri, sehingga konsumen memilih sistem operasi yang akan mereka pakai tergantung selera. Selain sistem operasi, *processor* yang digunakan juga turut mendukung *performance* sebuah *smartphone*, *processor* yang digunakan disini adalah banyaknya otak (*core*) yang digunakan dalam produk, terdapat empat jenis *processor* yakni *single core*, *dual core*, *quad core*, dan *octa core*, semakin banyaknya *core* yang digunakan

menjadikan *performance smartphone* semakin baik.

## 2. Kamera

Fenomena *selfie* di kalangan remaja maupun dewasa, menjadikan kamera sebagai patokan utama seseorang dalam memilih produk, semakin tinggi resolusi kamera yang digunakan (baik kamera utama maupun kamera depan), tingkat popularitasnya pun semakin tinggi, apalagi ditambah dengan fitur *flash* menjadikan produk tersebut digandrungi masyarakat.

## 3. Kapasitas memori yang digunakan (*Random Operation Memory* / ROM)

Kapasitas memori internal (ROM) dalam menyimpan *file* menjadi alasan konsumen dalam memilih produk, memori internal pada sebuah *smartphone* berkisar dari 4 *GigaByte* (GB), 8 GB, 16 GB, 32 GB, 64 GB, hingga 128 GB. Semakin tinggi memori internal yang tersedia, maka *smartphone* tersebut akan menjadi idaman bagi para konsumen.

Selain memori internal, di dalam *smartphone* juga terdapat *chip* memori yang mengatur kinerja dalam melaksanakan aplikasi, yaitu *Random Access Memory* (RAM), RAM adalah sebuah perangkat keras yang berfungsi menyimpan data dan instruksi program yang akan dan sudah dieksekusi oleh *processor*. Semakin banyak *free space* dalam RAM tersebut maka semakin cepat kinerja *processor* menjalankan sebuah aplikasi. RAM untuk *smartphone* berkisar antara 512 *MegaByte* (MB), 768 MB, 1 GB, 1,5 GB, 2 GB, 3 GB, hingga 4 GB, semakin besar RAM yang tersedia, maka peminat produk tersebut akan semakin banyak. Selain kedua memori tersebut penambahan memori eksternal juga menjadi salah satu yang dapat meningkatkan minat masyarakat.

#### 4. Jaringan

Perkembangan teknologi nirkabel yang sangat pesat, dimulai dari teknologi 2G, 3G, lalu sekarang 4G, diiringi dengan perkembangan spesifikasi yang lain menyebabkan *smartphone* menjadi perangkat yang sudah menjadi keharusan dalam kehidupan sehari-hari. Semakin tinggi jaringan yang digunakan maka kemampuan mengakses internet pun akan semakin tinggi pula, sehingga semakin menjadi incaran konsumen. Selain fitur mengakses internet, fitur Dual SIM pun menjadi fitur yang turut berkembang belakangan ini, kemampuan untuk mengakses dua kartu SIM sekaligus dapat menarik konsumen untuk menggunakan produk tersebut.

Selain fitur dual SIM, jenis ukuran SIM juga mengindikasikan tingkat kekinian sebuah produk, jika yang digunakan adalah Nano SIM, maka produk tersebut adalah produk terbaru yang diproduksi, begitu pula selanjutnya dengan Micro SIM, dan produk keluaran lama dengan Mini SIM.

#### 5. Kapasitas Baterai

Semakin tinggi spesifikasi yang dimiliki oleh sebuah produk, maka tingkat konsumsi baterainya pun semakin besar, hal ini yang kemudian menjadikan kapasitas baterai dalam patokan seorang konsumen memilih sebuah produk.

#### 6. Ukuran Layar

Untuk ukuran layar, semakin besar layar yang digunakan maka semakin memudahkan *user* dalam menggunakan produk. Namun, untuk ukuran layar sendiri semuanya tergantung selera masing-masing konsumen, ada yang tidak mau menggunakan layar yang terlalu besar ataupun yang terlalu

kecil, tetapi ukuran layar merupakan salah satu prioritas dalam memilih sebuah produk.

#### 7. Harga

Ini adalah hal paling utama yang dipertimbangkan konsumen dalam memilih produk. Produk dengan spesifikasi yang mumpuni, namun tidak terlalu mahal harganya menjadi pilihan utama. Namun, ada produk dengan harga relatif tinggi, namun spesifikasi menengah menjadi incaran, biasanya hal ini disebabkan oleh faktor *brand* yang menentukan tingkat sosial pengguna.

#### 8. Tingkat kepercayaan terhadap *brand*

Tingkat kepercayaan terhadap *brand* menyebabkan suatu produk digemari masyarakat, pada *brand-brand* produk yang sudah terkenal seperti Samsung, Apple, atau Sony, cenderung dipilih masyarakat karena sudah dipercaya memiliki produk-produk dengan kualitas tinggi.

Berdasarkan penjelasan di atas, maka dapat diambil beberapa peubah yang kemudian menjadi dasar penggerombolan produk-produk *smartphone* sehingga dapat dilihat karakteristik masing-masing produk yang beredar di masyarakat.

# BAB III

## PEMBAHASAN

*LCCA* digunakan untuk mengklasifikasikan individu atau objek ke dalam gerombol/kelas yang berbeda berdasarkan pada respon untuk sekumpulan peubah yang diamati. Tipe peubah-peubah yang mendasari *LCCA* dapat bertipe kategorik seluruhnya, numerik seluruhnya maupun bertipe campuran (numerik dan kategorik). Parameter model yang digunakan ditaksir menggunakan *maximum likelihood* dengan algoritma EM (Moustaki dan Papageorgiou, 2004).

Pada bab ini akan diuraikan metode *LCCA* berdasarkan peubah bertipe campuran dan implementasinya pada kasus pengelompokan produk *smartphone* menggunakan *software* LatentGOLD 4.0.

### 3.1 Model *Latent Class Cluster Analysis* Berdasarkan Peubah Bertipe Campuran

Model kelas laten yang digunakan oleh Moustaki dan Papageorgiou (2004) yaitu sebagai berikut:

**Definisi 3.1.1.** Misalkan  $(x_1, x_2, \dots, x_i)$  menunjukkan vektor dari  $i$  peubah teramati di mana masing-masing peubah memiliki distribusi bersyarat dalam keluarga eksponensial seperti *Bernoulli*, Multinomial, dan Normal. Peubah  $X$  dapat memiliki jenis yang sama, baik diskrit atau kontinu ataupun kombinasi dari kedua jenis tersebut. Misalkan  $x_{ih}$  adalah nilai objek ke  $h$  untuk peubah ke- $i$ , ( $h = 1, 2, \dots, N$ ).

Vektor baris  $x'_h = (x_{1h}, x_{2h}, \dots, x_{ih})$  adalah respon objek ke  $h$ .

Pada model-model kelas laten diasumsikan bahwa pada suatu ruang faktor terdapat  $J$  gerombol. Peluang awal (*prior*) untuk data pengamatan  $x$  pada setiap gerombol dinotasikan oleh  $\eta_j$ . Fungsi peluang bersama (*joint distribution*) dari peubah-peubah yang teramati menurut Moustaki dan Papageorgiou (2004) didefinisikan oleh :

$$f(x_h) = \sum_{j=1}^J \eta_j g(x_h|j) \quad (3.1)$$

dimana:

1.  $\eta_j$  adalah peluang awal (*prior*) untuk peubah  $x$  yang teramati pada gerombol ke- $j$  ( $\eta_j = \frac{n_j}{N}$ ),
2.  $g(x_h|j)$  adalah fungsi peluang untuk vektor peubah  $x$  pada objek  $h$  dari gerombol ke- $j$ .

Selanjutnya akan dibahas model kelas laten hanya untuk tipe peubah biner, nominal, ordinal, dan normal. Peluang bersyarat  $h(j|x_h)$  adalah peluang bahwa sebuah objek akan masuk ke dalam gerombol ke- $j$ , jika diketahui distribusi peluang dari objek tersebut ( $f(x_h)$ ), didapat dengan menggunakan teorema bayes, yaitu sebagai berikut:

$$\begin{aligned} h(j|x_h) &= \frac{\eta_j g(x_h|j)}{\sum_{j=1}^J \eta_j g(x_h|j)} \\ &= \frac{\eta_j g(x_h|j)}{f(x_h)} \end{aligned} \quad (3.2)$$

Fungsi  $g(x_h|j)$  adalah fungsi peluang bersama dari setiap model kelas laten pada setiap tipe data respon, dimana persamaan  $g(x_h|j)$  adalah sebagai

berikut:

$$\begin{aligned}
g(x_h|j) &= \prod_{i=1}^P g(x_{ih}|j), \quad j = 1, 2, \dots, J \\
&= \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S \pi_{ij(s)}^{x_{ih(s)}} \right] \\
&\quad \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right]
\end{aligned}$$

dengan  $B$  adalah banyaknya peubah teramati yang bertipe biner,  $L$  adalah banyaknya peubah teramati yang bertipe nominal,  $O$  adalah banyaknya peubah teramati yang bertipe ordinal dan  $Q$  adalah banyaknya peubah teramati yang bertipe kontinu.

Model campuran bersama dari peubah-peubah yang diamati untuk  $LC-CA$  model ini, dapat dituliskan sebagai berikut:

$$\begin{aligned}
f(x_h) &= \sum_{j=1}^J \eta_j \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \right] \right] \\
&\quad \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right]
\end{aligned}$$

Metode yang paling sering digunakan untuk penaksiran parameter pada  $h(j|x_h)$  adalah dengan menggunakan metode kemungkinan maksimum (*maximum likelihood*). Solusi dari suatu fungsi *likelihood* seringkali tidak dapat diperoleh dengan menggunakan turunan (*derivatif*), namun seringkali penyelesaiannya diperoleh melalui metode iterasi (numerik).

## 3.2 Penaksiran Parameter

Salah satu metode untuk penaksiran parameter dalam model kelas laten adalah maksimum *likelihood* (ML). Fungsi *log-likelihood* yang disyaratkan pada pendekatan ML dapat diperoleh dari fungsi kepadatan peluang yang mendefinisikan model. Algoritma EM (*Expectation-Maximization Algorithm*) merupakan salah satu algoritma untuk mendapatkan solusi dari fungsi *likelihood*, dimana solusi tersebut merupakan taksiran untuk parameter yang memaksimumkan fungsi *log-likelihood*.

Fungsi *likelihood* untuk sample acak berukuran  $N$  berdasarkan fungsi peluang vektor peubah- $x$  pada objek- $h$  ( $f(x_h)$ ) adalah:

$$\begin{aligned} L &= \prod_{h=1}^N f(x_h) \\ &= \prod_{h=1}^N \sum_{j=1}^J \eta_j g(x_h|j) \end{aligned} \quad (3.3)$$

Fungsi *likelihood* untuk model kelas laten dengan peubah- $x$  bertipe campuran biner, nominal, ordinal, dan normal, dapat dituliskan menjadi:

$$\begin{aligned} L &= \prod_{h=1}^N \sum_{j=1}^J \eta_j g(x_h|j) \\ &= \prod_{h=1}^N \sum_{j=1}^J \eta_j \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \right] \right] \\ &\quad \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ -\sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right] \end{aligned}$$

Fungsi *log-likelihood* berdasarkan fungsi *likelihood* di atas diberikan oleh:

$$\begin{aligned}
\ln L &= \ln \prod_{h=1}^N \sum_{j=1}^J \eta_j \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \right] \right. \\
&\quad \left. \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right] \right] \\
&= \sum_{h=1}^N \ln \sum_{j=1}^J \eta_j \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \right] \right. \\
&\quad \left. \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right] \right]
\end{aligned}$$

Menurut Nainggolan (2008), tahapan algoritma EM ini dapat dibagi menjadi beberapa tahapan. Terlebih dahulu adalah penentuan nilai parameter awal  $\theta^{(0)} = (\eta^{(0)}, \pi^{(0)}, \pi_s^{(0)}, \gamma_s - \gamma_{s-1}^{(0)}, \mu^{(0)}, \text{ dan } \sigma^{(0)})$ . Misalkan urutan pengulangan dinotasikan dengan  $t$ . Proses penaksiran pada algoritma EM dimulai dengan iterasi dari EM. Setiap lingkaran proses pada algoritma EM terdiri dari dua langkah, yaitu pada langkah ekspektasi dan pemaksimuman dengan tahapan:

1. Mendefinisikan nilai awal  $\theta^{(0)} = (\eta^{(0)}, \pi^{(0)}, \pi_s^{(0)}, \gamma_s - \gamma_{s-1}^{(0)}, \mu^{(0)}, \text{ dan } \sigma^{(0)})$ ,
2. Menghitung nilai dari  $f(x_h)^{(0)}$ , yaitu:

$$\begin{aligned}
f(x_h)^{(0)} &= \sum_{j=1}^J \eta_j^{(0)} \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}^{(0)}} (1 - \pi_{ij})^{1-x_{ih}^{(0)}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}^{(0)}} \right] \right] \\
&\quad \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}^{(0)}} \right] \right] \\
&\quad \left[ \left[ \frac{1}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(0)}} \exp \left[ - \sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right]^{(0)}
\end{aligned}$$

3. Tahapan E: Menghitung  $\hat{h}(j|x_h)^{(t)}$ , dengan  $h = 1, 2, \dots, N$ ;  $j = 1, 2, \dots, J$ , dimana  $\hat{h}(j|x_h)^{(t)}$  adalah peluang bersyarat objek akan masuk ke dalam kelas ke- $j$  jika diketahui berdistribusi  $f(x_h)$ , yang dirumuskan sebagai berikut:

$$\hat{h}(j|x_h)^{(t)} = \frac{\hat{\eta}_j^{(t-1)} \hat{g}(x_h|j)^{(t-1)}}{\sum_{j=1}^J \hat{\eta}_j^{(t-1)} \hat{g}(x_h|j)^{(t-1)}} \quad (3.4)$$

4. Tahapan M: menghitung nilai maksimum *likelihood* dengan menggunakan nilai parameter yang diasumsikan sama dengan nilai tahapan ekspektasi, sehingga diperoleh dugaan parameter yang baru, yaitu sebagai berikut:

- (a) Taksiran peluang awal setiap kelas ke- $j$  adalah:

$$\hat{\eta}_j^{(t)} = \frac{\sum_{h=1}^N \hat{h}(j|x_h)^{(t)}}{N} \quad (3.5)$$

- (b) Taksiran peluang bersyarat bahwa  $x_{ih} = 1$  pada kelas ke- $j$  untuk peubah biner adalah:

$$\hat{\pi}_{ij}^{(t)} = \frac{\sum_{h=1}^N x_{ih} \hat{h}(j|x_h)^{(t)}}{N \hat{\eta}_j^{(t)}} \quad (3.6)$$

- (c) Taksiran peluang bersyarat bahwa  $x_{ih} = s$  pada kelas ke- $j$  untuk peubah nominal adalah:

$$\hat{\pi}_{ij(s)}^{(t)} = \frac{\sum_{h=1}^N x_{ih(s)} \hat{h}(j|x_h)^{(t)}}{N \hat{\eta}_j^{(t)}} \quad (3.7)$$

- (d) Taksiran kumulatif peluang bersyarat dengan  $x_{ih} \leq s$  pada kelas ke- $j$

untuk peubah ordinal adalah:

$$\hat{\gamma}_{ij(s)}^{(t)} = \hat{\gamma}_{ij(s-1)}^{(t)} + \frac{\sum_{h=1}^N x_{ih(s)} \hat{h}(j|x_h)^{(t)}}{N \hat{\eta}_j^{(t)}} \quad (3.8)$$

(e) Taksiran rataaan untuk peubah kontinu kelas ke- $j$  adalah:

$$\hat{\mu}_{ij}^{(t)} = \frac{\sum_{h=1}^N x_{ih} \hat{h}(j|x_h)^{(t)}}{N \hat{\eta}_j^{(t)}} \quad (3.9)$$

(f) Taksiran *varians* (diasumsikan konstan/*fixed*) setiap kelas ke- $j$  adalah:

$$\hat{\sigma}_i^2{}^{(t)} = \frac{\sum_{h=1}^N \sum_{j=1}^J (x_{ih} - \hat{\mu}_{ij})^2{}^{(t)} \hat{h}(j|x_h)^{(t)}}{\sum_{h=1}^N \sum_{j=1}^J \hat{h}(j|x_h)^{(t)}} \quad (3.10)$$

(g) Taksiran untuk peluang posterior  $h(j|x_h)$  adalah:

$$\hat{h}(j|x_h)^{(t+1)} = \frac{\hat{\eta}_j^{(t)} \hat{g}(x_h|j)^{(t)}}{\hat{f}(x_h)^{(t)}} \quad (3.11)$$

5. Ulangi tahap 2,3, dan 4 sampai konvergen.

Untuk pembuktian persamaan 3.5 hingga 3.11 dapat dilihat pada lampiran 1.

Pada tahap awal pengelompokkan objek ke dalam gerombol dapat dilakukan secara acak atau menggunakan hierarki atau analisis gerombol *K-Means*. Pemilihan nilai awal yang berbeda-beda perlu dilakukan sebelum memilih model akhir, dengan tujuan untuk menghindari kekonvergensi fungsi *likelihood* menjadi maksimum lokal.

Klasifikasi akhir dari objek ke dalam gerombol yang telah teridentifikasi didasarkan pada taksiran peluang posterior ( $h(j|\hat{x}_h)$ ). Sebuah objek memiliki kemungkinan lebih tinggi dialokasikan ke gerombol- $j$  daripada gerombol- $k$ , jika  $h(j|x_h) > h(k|x_h)$ . Kelebihan jika distribusi  $g(x_{ih}|j)$  berasal dari keluarga eksponensial adalah sebaran peluang posteriornya merupakan fungsi linier dari  $x$  (Bartholomew dan Knott, 1999, dalam Moustaki dan Papageorgiou, 2004).

### 3.3 Pemilihan Model

Berdasarkan model yang dihasilkan dari analisis, maka akan dipilih model terbaik yang mampu menjelaskan data. *Bayesian Information Criterion* (BIC) adalah sebuah kriteria statistik untuk memilih model. Model dengan nilai BIC yang lebih kecil dipilih sebagai model yang terbaik, karena nilai BIC yang lebih kecil menunjukkan bahwa model yang dihasilkan lebih bisa menjelaskan variasi dari data. Nilai untuk BIC dapat diperoleh menggunakan rumus (Vermunt dan Magidson, 2004):

$$BIC = -2l(\hat{\theta}) + \ln(N)q \quad (3.12)$$

dimana:

$N$  adalah banyaknya objek pengamatan,

$q$  adalah banyaknya parameter,

$l(\hat{\theta})$  adalah nilai maksimum fungsi *log-likelihood* dari suatu model yang ditaksir.

Misalkan pada suatu kasus pengerombolan yang terdiri dari 52 objek pengamatan diperoleh hasil:

Tabel 3.1: Contoh Perhitungan BIC

	$l(\hat{\theta})$	$q$	BIC
1 Gerombol	-1151,32	8	2334,249
2 Gerombol	-950,819	17	1968,801
3 Gerombol	-929,146	26	1961,012
4 Gerombol	-908,637	35	1955,563
5 Gerombol	-891,071	44	1955,994

Dari perhitungan tersebut, untuk menentukan banyaknya gerombol yang terbentuk dipilih berdasarkan nilai BIC terkecil yang diperoleh. Nilai BIC terkecil didapatkan pada saat banyaknya gerombol yang terbentuk sebanyak 4, dengan nilai BIC sebesar 1955,563. Dengan demikian, banyaknya gerombol yang terbentuk dari kasus ini adalah sebanyak 4 gerombol.

## 3.4 Data

### 3.4.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari <http://www.hariangadget.com//> dan <http://www.gsmarena.com//>. Untuk peubah harga diperoleh dari <http://www.hariangadget.com//>, sedangkan untuk peubah-peubah lainnya diperoleh dari <http://www.gsmarena.com//>, peubah-peubah teramati inilah yang kemudian dijadikan dasar penggerombolan pada *Latent Class Cluster Analysis*.

Objek (banyaknya produk *smartphone*) yang dianalisis berjumlah 194 produk, yang berasal dari 9 produsen: Samsung, Lenovo, Sony, LG, Nokia, ASUS, Apple, BlackBerry dan Microsoft. Dari 194 produk *smartphone* tersebut, diantaranya 29,38% dari produsen Samsung, 17,53% dari Lenovo, 13,40% dari Sony,

12,89% dari LG, 10,82% dari Nokia, 6,19% dari ASUS, 5,67% dari Apple, 2,57% dari Blackberry, dan 1,54% dari Microsoft.

### 3.4.2 Identifikasi Peubah

Dalam penelitian ini, digunakan 15 peubah teramati, yaitu sebanyak 3 peubah bertipe biner, 2 peubah bertipe nominal, 8 peubah bertipe ordinal, dan 2 peubah bertipe kontinu (berdistribusi normal). Untuk penjelasan masing-masing peubah teramati dapat dilihat di tabel dibawah ini.

Tabel 3.2: Peubah Teramati yang Digunakan

No	Peubah	Definisi	Tipe	Respon	Parameter
1	$X_1$	Jaringan yang digunakan	Ordinal	1 : EDGE 2 : HSPA 3 : LTE	$\gamma_{1j(1)}$ $\gamma_{1j(2)}$ $\gamma_{1j(3)}$
2	$X_2$	Kemampuan Dual SIM	Biner	1 : Ya 0 : Tidak	$\pi_{2j}$ $1 - \pi_{2j}$
3	$X_3$	Jenis SIM yang digunakan	Nominal	1 : Mini SIM 2 : Micro SIM 3 : Nano SIM	$\pi_{3j(1)}$ $\pi_{3j(2)}$ $\pi_{3j(3)}$
4	$X_4$	Ukuran Layar (inch)	Ordinal	1 : < 4,00 inch 2 : 4,00 - 4,99 inch 3 : 5,00 - 5,99 inch 4 : 6,00 - 6,99 inch 5 : $\geq$ 7,00 inch	$\gamma_{4j(1)}$ $\gamma_{4j(2)}$ $\gamma_{4j(3)}$ $\gamma_{4j(4)}$ $\gamma_{4j(5)}$
5	$X_5$	Jenis OS yang digunakan	Nominal	1 : Android 2 : Windows 3 : iOS 3 : BlackBerry	$\pi_{5j(1)}$ $\pi_{5j(2)}$ $\pi_{5j(3)}$ $\pi_{5j(4)}$

Tabel 3.3: Peubah Teramati yang Digunakan (Lanjutan)

No	Peubah	Definisi	Tipe	Respon	Parameter
6	$X_6$	Tipe <i>core</i> yang digunakan	Ordinal	1 : <i>Single Core</i> 2 : <i>Dual Core</i> 3 : <i>Quad Core</i> 4 : <i>Octa Core</i>	$\gamma_{6j(1)}$ $\gamma_{6j(2)}$ $\gamma_{6j(3)}$ $\gamma_{6j(4)}$
7	$X_7$	Fasilitas Memori Eksternal	Biner	1 : Ya 0 : Tidak	$\pi_{7j}$ $1 - \pi_{7j}$
8	$X_8$	Memori Internal yang Tersedia (ROM)	Ordinal	1 : 4 GB 2 : 8 GB 3 : 16 GB 4 : 32 GB 5 : 64 GB 6 : 128 GB	$\gamma_{8j(1)}$ $\gamma_{8j(2)}$ $\gamma_{8j(3)}$ $\gamma_{8j(4)}$ $\gamma_{8j(5)}$ $\gamma_{8j(6)}$
9	$X_9$	RAM yang Tersedia	Ordinal	1 : 512 MB 2 : 768 MB 3 : 1 GB 4 : 1,5 GB 5 : 2 GB 6 : 3 GB 7 : 4 GB	$\gamma_{9j(1)}$ $\gamma_{9j(2)}$ $\gamma_{9j(3)}$ $\gamma_{9j(4)}$ $\gamma_{9j(5)}$ $\gamma_{9j(6)}$ $\gamma_{9j(7)}$
10	$X_{10}$	Besarnya Kamera Utama (MP)	Ordinal	1 : <5,00 MP 2 : 5,00-7,99 MP 3 : 8,00-12,99 MP 4 : $\geq 13,00$ MP	$\gamma_{10j(1)}$ $\gamma_{10j(2)}$ $\gamma_{10j(3)}$ $\gamma_{10j(4)}$
11	$X_{11}$	Fasilitas <i>Flash</i>	Biner	1 : Ya 0 : Tidak	$\pi_{11j}$ $1 - \pi_{11j}$
12	$X_{12}$	Besarnya Kamera Depan (MP)	Ordinal	1 : Tidak ada 2 : 0,01-1,99 MP 3 : 2,00-4,99 MP 4 : $\geq 5,00$ MP	$\gamma_{12j(1)}$ $\gamma_{12j(2)}$ $\gamma_{12j(3)}$ $\gamma_{12j(4)}$
13	$X_{13}$	Harga* (Juta Rupiah)	Kontinu	Mean Varians	$\mu_{13j}$ $\sigma_{13j}^2$
14	$X_{14}$	Tingkat Popularitas** (Juta Rupiah)	Kontinu	Mean Varians	$\mu_{14j}$ $\sigma_{14j}^2$
15	$X_{15}$	Kapasitas Baterai (mAh)	Ordinal	1 : $\leq 2000$ mAh 2 : 2001-3000 mAh 3 : 3001-4000 mAh 4 : $> 4000$ mAh	$\gamma_{15j(1)}$ $\gamma_{15j(2)}$ $\gamma_{15j(3)}$ $\gamma_{15j(4)}$

Keterangan:

\* : Harga pada Mei 2015, diambil dari *www.hariangadget.com*,

\*\* : Tingkat popularitas berdasarkan survey pihak *www.gsmarena.com*.

Semua peubah kontinu pada penelitian ini yaitu peubah harga dan tingkat popularitas diasumsikan berdistribusi normal. Untuk melakukan uji normalitas terhadap peubah-peubah tersebut digunakan uji *Saphiro-Wilk*.

Dari analisis perhitungan pada tabel 3.4 terlihat bahwa asumsi normalitas belum terpenuhi pada kedua peubah kontinu tersebut, atau dengan kata lain peubah-peubah tersebut tidak berdistribusi normal ( $p\text{-value} < 0,05$ ). Dengan demikian, dilakukan transformasi data pada peubah harga dan popularitas. Bentuk transformasi yang digunakan adalah transformasi *logaritma-natural*.

Tabel 3.4: Uji Normalitas Data Sebelum Ditransformasi

Peubah	Definisi	$p\text{-value}$
$X_{13}$	Harga (Juta Rupiah)	0,0000
$X_{14}$	Tingkat Popularitas (Persen)	0,0000

Setelah dilakukan transformasi *logaritma-natural*, kedua peubah tersebut sudah memenuhi asumsi normalitas ( $p\text{-value} > 0,05$ ), seperti ditunjukkan dalam tabel 3.5 berikut ini.

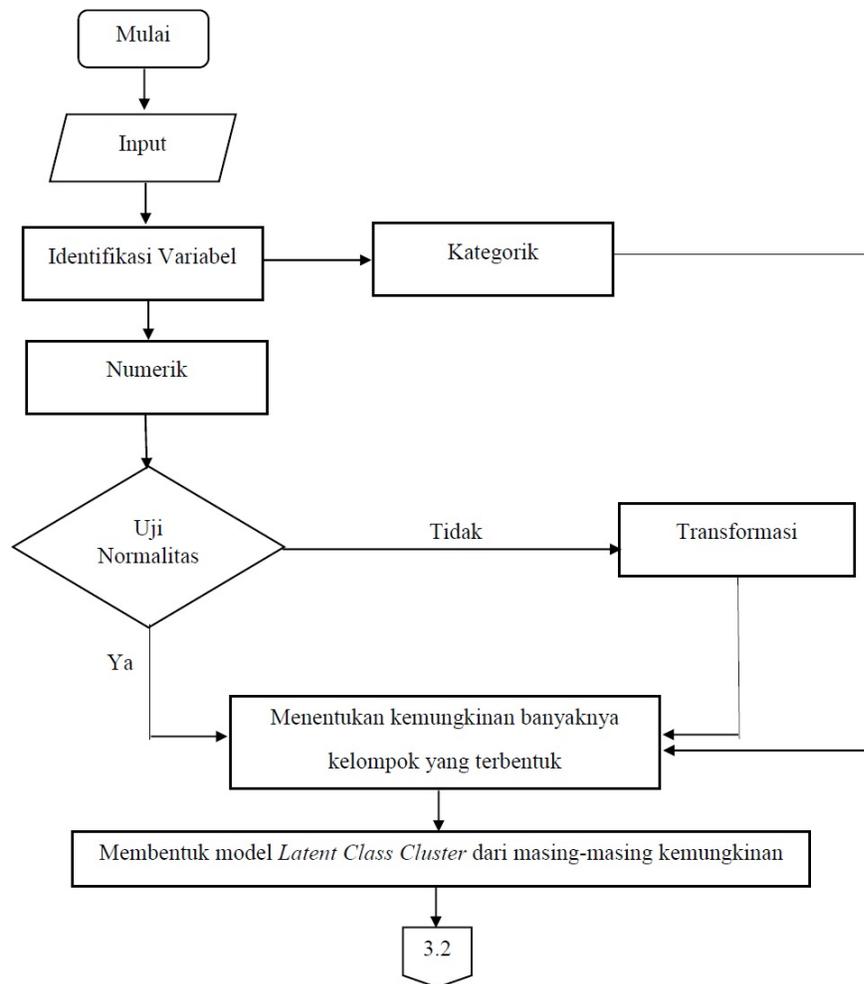
Tabel 3.5: Uji Normalitas Data Setelah Ditransformasi

Peubah	Definisi	$p\text{-value}$
$\ln X_{13}$	Nilai ln dari Harga (Juta Rupiah)	0,255
$\ln X_{14}$	Nilai ln dari Tingkat Popularitas (Persen)	0,262

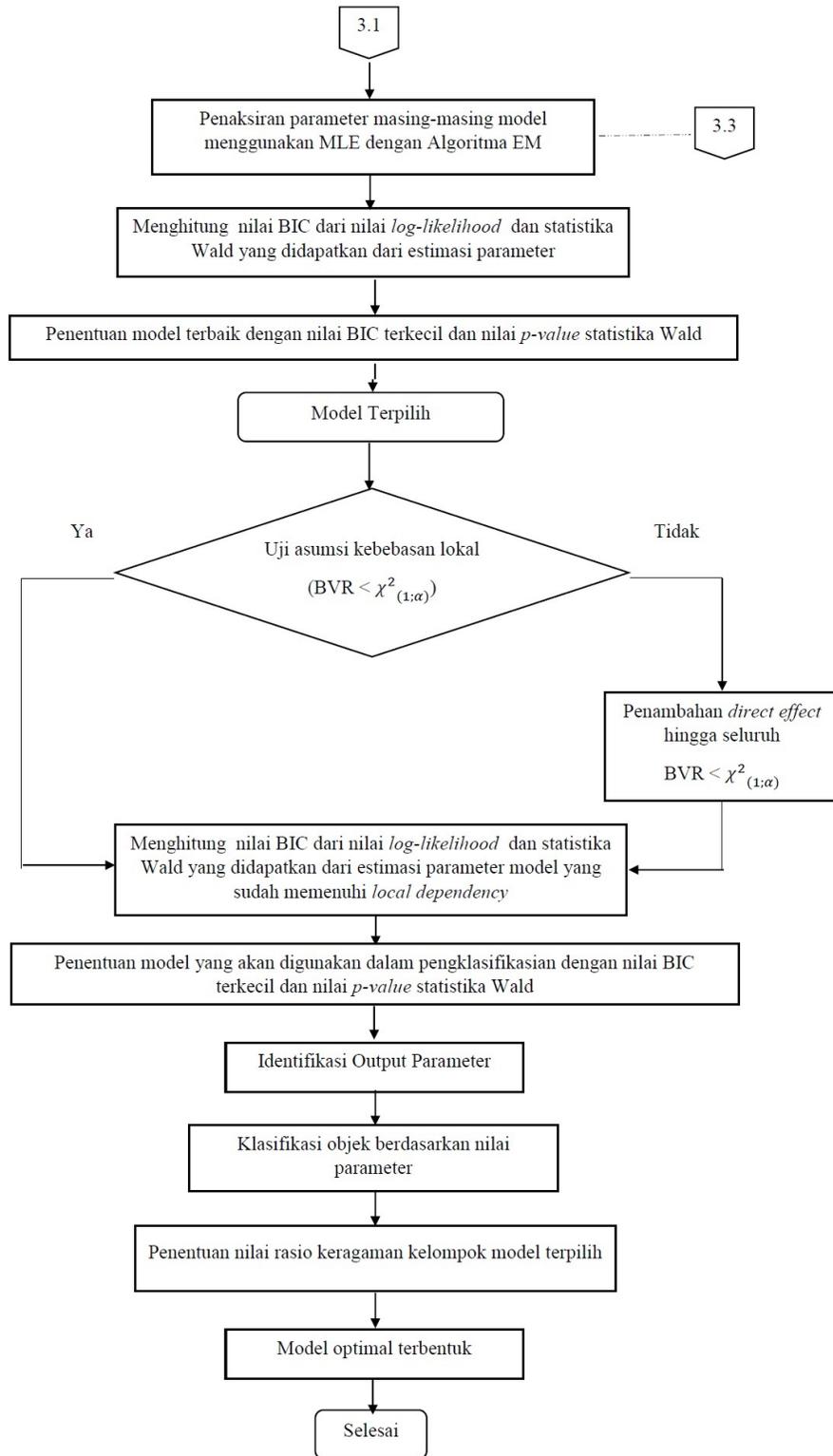
Dengan demikian, untuk proses penggerombolan selanjutnya akan didasarkan 13 peubah bertipe kategorik, dan 2 peubah normal yang telah ditransformasi.

### 3.5 Metode Analisis Data

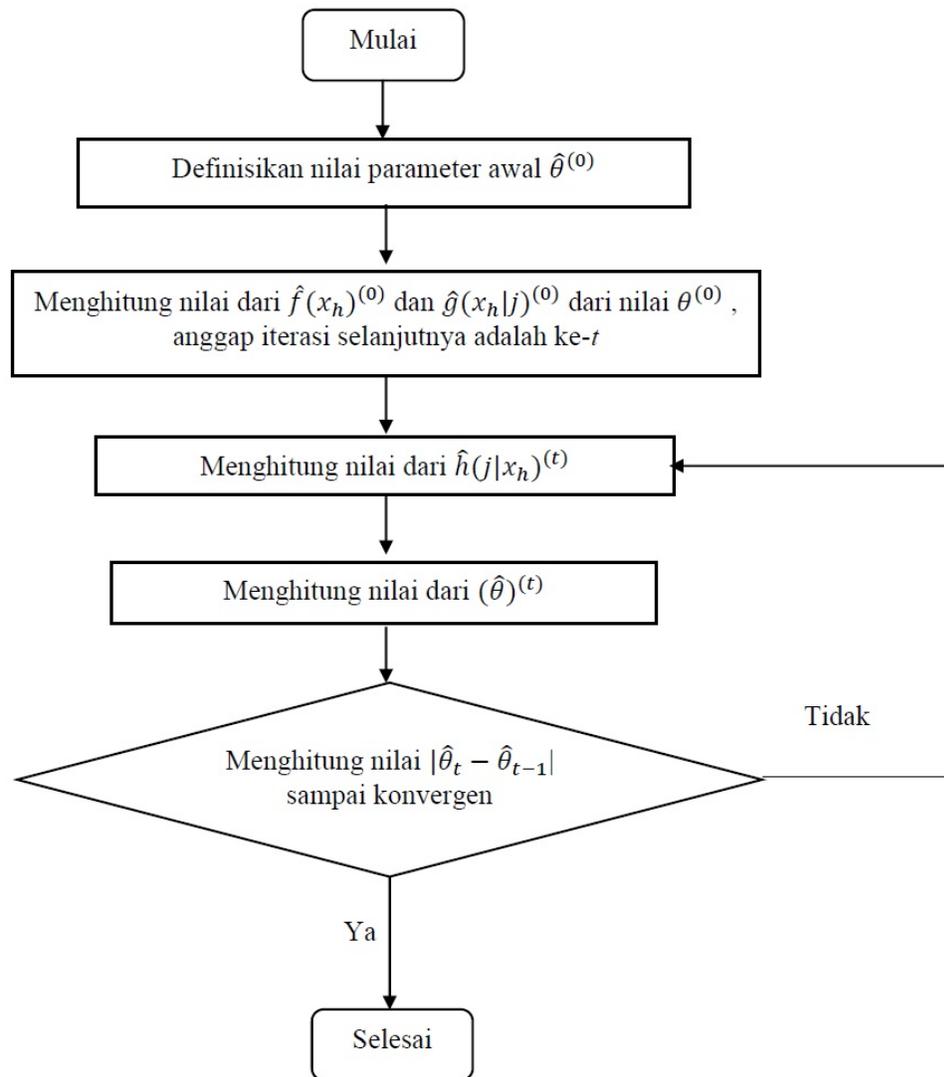
Metode analisis data yang digunakan dalam penelitian ini adalah *LCCA*. Tujuan dari metode ini adalah mengelompokkan objek-objek ke dalam kelas-kelas laten berdasarkan peubah-peubah campuran yang diamati. Adapun langkah-langkah dari penggunaan *LCCA* dapat dijelaskan dalam diagram alir sebagai berikut:



Gambar 3.1: Diagram Alir *Latent Class Cluster Analysis*



Gambar 3.2: Diagram Alir *Latent Class Cluster Analysis* (Lanjutan)



Gambar 3.3: Diagram Alir Algoritma EM

Adapun penjelasan dari setiap tahapan metode ini adalah sebagai berikut:

1. Melakukan input data terhadap peubah kategorik dan numerik yang digunakan, lalu mengidentifikasi skala pengukuran pada setiap peubah, dan melakukan uji normalitas terhadap data kontinu yang digunakan, jika belum terpenuhi asumsi normalitasnya, maka dilakukan transformasi.

2. Menentukan kemungkinan banyaknya gerombol yang akan dibuat, biasanya dibatasi hingga 5 gerombol, tergantung keinginan dan kebutuhan peneliti (Budiati, 2014), dan membentuk model *LCCA* dari masing-masing kemungkinan gerombol yang terbentuk.
3. Melakukan penaksiran parameter menggunakan metode maksimum *likelihood* yang dilakukan dengan iterasi algoritma EM.
4. Mencari nilai BIC dari masing-masing model gerombol yang terbentuk dengan berdasarkan nilai *log-likelihood* serta uji kelayakan model dengan menggunakan uji statistik *Wald*. Jika *p-value* dari statistik *Wald* sudah kurang dari 0,05, maka model dianggap sudah layak. Model sudah layak dan memiliki nilai BIC terkecil merupakan model terpilih yang akan digunakan.
5. Model terpilih perlu dilakukan uji asumsi kebebasan lokal. Asumsi kebebasan lokal terpenuhi jika nilai  $BVR < \chi^2_{(1,\alpha)}$ . Jika asumsi kebebasan lokal belum terpenuhi maka perlu dilakukan penambahan *direct effect* kedalam model hingga seluruh nilai BVR antara pasangan peubah mencapai kurang dari  $\chi^2_{(1,\alpha)}$  atau kurang dari 6,63 (Vermunt dan Magidson, 2004).
6. Model yang telah layak dan memenuhi asumsi kebebasan lokal, perlu dilihat kembali nilai BIC dan statistik *Wald*-nya, jika BIC sudah terkecil dan statistik *Wald* terpenuhi, maka dilakukan pengklasifikasian objek (*cluster membership*) didasarkan pada nilai taksiran parameter model yang telah dipilih tersebut.
7. Menentukan nilai keragaman gerombol model optimal yang terbentuk dengan membagi antara nilai jumlah kuadrat dalam gerombol (SSW) dengan

nilai jumlah kuadrat antar gerombol (SSB). Semakin kecil rasio keragaman gerombol, model yang dibuat semakin baik.

## 3.6 Hasil Analisis Data

Pada penelitian ini akan dibahas mengenai bagaimana penggerombolan 194 produk *smartphone* dari berbagai produsen, dan akan dianalisis bagaimana karakteristik masing-masing gerombol optimal yang telah dibuat.

### 3.6.1 Penaksiran Peluang Awal Kelas Laten

Penaksiran parameter dalam model kelas laten untuk peubah bertipe campuran biner, nominal, ordinal, dan normal digunakan untuk mengklasifikasikan objek ke dalam kelas laten. Peluang awal (*prior*) kelas laten adalah nilai peluang banyaknya objek dalam suatu gerombol terhadap keseluruhan gerombol. Penaksiran peluang awal dilakukan dengan menggunakan metode maksimum *likelihood*. Untuk memperoleh solusi dari metode maksimum *likelihood* tersebut digunakan algoritma EM.

Penaksiran peluang awal ( $\eta_j$ ) dilakukan hingga 5 gerombol. Nilai taksiran untuk peluang awal diberikan pada Tabel 3.6.

Tabel 3.6: Peluang Awal Kelas Laten ( $\eta_j$ ) Untuk Setiap Kemungkinan Model

Model	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$	$\eta_5$
1 Gerombol	1,0000				
2 Gerombol	0,6256	0,3744			
3 Gerombol	0,5200	0,3445	0,1355		
4 Gerombol	0,3629	0,3090	0,2080	0,1201	
5 Gerombol	0,3151	0,2941	0,2122	0,1193	0,0593

Dari Tabel 3.6, dapat dilihat bahwa untuk model dengan 5 gerombol, taksiran peluang awal untuk gerombol 1 ( $\eta_1$ ) dan gerombol 2 ( $\eta_2$ ) masing-masing adalah 0.3151 dan 0.2941, artinya sebanyak 31.5% dari produk yang dianalisis akan masuk ke gerombol pertama dan 29.41% produk masuk ke gerombol 2. Interpretasi yang sama juga diberikan untuk taksiran peluang awal gerombol 3, 4 dan 5. Dari berbagai kemungkinan model yang disajikan pada Tabel 3.6, akan ditentukan model terbaik untuk penggerombolan 194 produk *smartphone* yang dianalisis.

### 3.6.2 Pemilihan Model Terbaik

Setelah dilakukan uji pada beberapa peubah laten dengan penaksiran yang disajikan dari 1 gerombol hingga 5 gerombol, kemudian menentukan penggerombolan yang cocok berdasarkan nilai BIC. Semakin kecil nilai BIC maka akan semakin baik model yang diperoleh. Nilai BIC pada setiap kelas laten dapat dilihat pada Tabel 3.7.

Tabel 3.7: Kriteria Untuk Pemilihan Model Terbaik

Model	Banyaknya Gerombol	<i>LogLikelihood</i>	BIC (LL)	Npar	<i>Class. Err</i>	<i>Wald</i>	<i>p-value</i>
Model 1	1 Gerombol	-3104,4713	6424,9247	41	0,0000	-	-
Model 2	2 Gerombol	-2682,1308	5890,8688	62	0,0196	9,7685	0,0018
Model 3	3 Gerombol	-2517,7565	5472,7452	83	0,0146	37,4790	7,3e-9
Model 4	4 Gerombol	-2365,3537	5278,5646	104	0,0221	24,6958	1,8e-5
Model 5	5 Gerombol	<b>-2261,7041</b>	<b>5181,8905</b>	125	0,0198	37,4457	<b>1,5e-7</b>

Keterangan :

Npar : Banyaknya parameter pada model,

*Class. Err* : Nilai kesalahan klasifikasi dari model yang terbentuk.

Dari Tabel 3.7 diperoleh nilai BIC terkecil yaitu 5181,8905 yang terdapat pada model 5. Model ini terdiri dari 5 gerombol atau 5 kelas laten dengan jumlah parameter sebanyak 125 dan nilai *log-likelihood* -2261,7041. Grafik nilai BIC dari Model 1 hingga Model 5 dapat dilihat pada Gambar 3.4. Semakin banyak gerombol yang terbentuk, maka nilai BIC nya akan semakin kecil.



Gambar 3.4: Nilai BIC Model 1 Gerombol Hingga 5 Gerombol

Menurut Vermunt dan Magidson (2002), kriteria untuk pemilihan model terbaik adalah model yang memiliki nilai BIC terkecil, dan memiliki *p-value* statistik *Wald* yang kurang dari 0,05. Berdasarkan Tabel 3.7, model 5 gerombol ini terpilih sebagai model terbaik, karena memiliki nilai BIC terkecil (5181,8905), dan *p-value* yang kurang dari 0,05 (yaitu sebesar  $1.5 \times 10^{-7}$ ).

Model 5 ini memperlihatkan bahwa dalam gerombol 1 terdapat 62 objek (31,51%), gerombol 2 ada 57 objek (29,41%), gerombol 3 ada 41 objek (21,22%), gerombol 4 ada 23 objek (11,93%), dan gerombol 5 ada 11 objek (5,93%). Penaksiran parameter secara iterasi mencapai kekonvergenan pada nilai *log-likelihood*

sebesar -2261,7041. Berikut alokasi objek ke dalam setiap gerombol dapat dilihat di tabel 3.8.

Tabel 3.8: Alokasi Objek Pada Setiap Gerombol

Gerombol	1	2	3	4	5	Total
1	60,8396	0,3258	0,0000	0,0000	0,0726	61,2379
2	1,1235	55,2095	0,8125	0,0000	0,0048	57,1474
3	0,0000	0,9905	40,1875	0,0000	0,0000	41,1780
4	0,0315	0,0369	0,0000	23,0000	0,0000	23,0684
5	0,0054	0,4403	0,0000	0,0000	10,9226	11,3683
Total	62,0000	57,0000	41,0000	23,0000	11,0000	194,0000

### 3.6.3 Uji Asumsi Kebebasan Lokal Model Terpilih

Asumsi kebebasan lokal adalah asumsi dasar pada model *LCCA*. Ketidakcocokan model dalam *LCCA* disebabkan oleh pelanggaran asumsi ini. Menurut Budiati (2014), jika ditemukan adanya *local dependency*, cara untuk menghilangkannya adalah dengan menambah jumlah gerombol sehingga tidak ditemukan adanya *local dependency*.

Kriteria model yang paling cocok adalah jika nilai residual seluruh pasangan peubah kategorik memenuhi asumsi kebebasan lokal (yaitu memiliki  $BVR < 6,63$ ). Alternatif untuk menghilangkan *local dependency* selain menambah jumlah gerombol, yaitu dengan melibatkan *direct effect* ke dalam model dengan tujuan memperkecil nilai residual pasangan peubah kategorik hingga tercapai seluruh nilai  $BVR$  kurang dari 6,63 (Vermunt dan Magidson, 2004).

Pada model terpilih dengan nilai BIC terkecil didapatkan 5 gerombol optimal dalam penggerombolan 194 objek. Namun, pada model ini masih terdapat 5 pasangan peubah kategorik yang memiliki nilai  $BVR > 6,63$  atau tidak memenuhi asumsi kebebasan lokal (nilai seluruh  $BVR$  dapat dilihat pada lampi-

ran 3). Dengan demikian, perlu ditambahkan *direct effect* ke dalam model 5 ini hingga seluruh BVR pasangan peubah kategorik kurang dari 6,63.

Untuk kasus ini, diperlukan penambahan *direct effect* selama 6 kali. Setelah ditambahkan 6 kali *direct effect*, didapatkan seluruh nilai BVR sudah memenuhi asumsi kebebasan lokal dengan selang BVR 0-6,4175 ( $BVR < 6,63$ ). Hasil penambahan *direct effect* ini dapat dilihat pada Tabel 3.9 dibawah ini.

Tabel 3.9: Kriteria Untuk Pemilihan Model Terbaik dengan Menggunakan *Direct Effect*

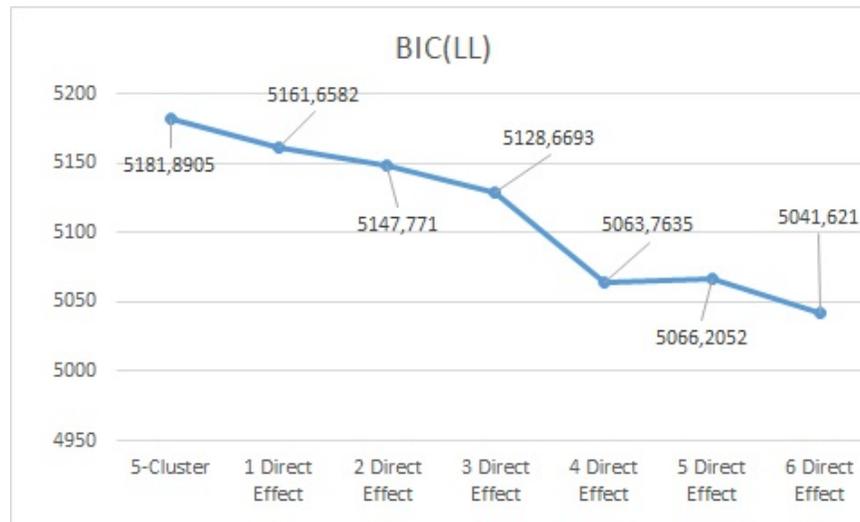
Model	<i>logLikelihood</i>	<i>BIC (LL)</i>	Npar	<i>Class Err</i>	BVR Maks.	<i>Wald</i>	<i>p-value</i>
5 Gerombol	-2261,7041	5181,8905	125	0,0198	8,9841	37,4457	1,5e-7
1 <i>d.e.</i>	-2248,9540	5161,6582	126	0,0168	7,7335	40,6841	3,1e-8
2 <i>d.e.</i>	-2239,3765	5147,7710	127	0,0167	7,9577	38,9516	7,1e-8
3 <i>d.e.</i>	-2227,1917	5128,6693	128	0,0163	7,8189	40,1465	4,0e-8
4 <i>d.e.</i>	-2192,1049	5063,7635	129	0,0192	7,8688	46,1623	2,3e-9
5 <i>d.e.</i>	-2185,4240	5066,2052	132	0,0212	7,7344	44,7246	4,5e-9
6 <i>d.e.</i>	<b>-2170,4980</b>	<b>5041,6210</b>	133	0,0205	<b>6,4175</b>	53,5305	<b>6,6e-11</b>

Keterangan :

Npar : Banyaknya parameter pada model,

*Class. Err* : Nilai kesalahan klasifikasi dari model yang terbentuk.

Dari Tabel 3.9 dapat dilihat bahwa model dengan 5 gerombol dan penambahan 6 *direct effect* memiliki nilai BIC terkecil yaitu 5041,6210. Model 5 gerombol dengan 6 *direct effect* ini memiliki jumlah parameter sebanyak 133 dan nilai *log-likelihood* -2170,4980. Gambar 3.5 menyajikan nilai BIC untuk model dengan 5 gerombol beserta penambahan *direct effect* sebanyak 1 hingga 6 kali.



Gambar 3.5: Nilai BIC dengan Penambahan *Direct Effect*

Kelayakan model juga dapat dilihat melalui *p-value* dari statistik *Wald*. Model yang dianggap layak adalah model dengan *p-value* kurang dari 0,05 (Ver-munt dan Magidson, 2004). Dari Tabel 3.9 dapat dilihat bahwa untuk model dengan 5 gerombol dan penambahan 6 *direct effect* memiliki *p-value* sebesar  $6.6 \times 10^{-11}$  (nilai ini kurang dari  $\alpha = 0.05$ ).

Dari berbagai penjelasan di atas didapatkan bahwa model 5 gerombol dengan 6 *direct effect* memiliki nilai BIC terkecil yaitu sebesar 5041,6210, dengan alokasi objek untuk gerombol pertama sebesar 33.86% (68 objek), gerombol kedua 32,64% (63 objek), gerombol ketiga 18,92% (35 objek), gerombol keempat 9,86% (19 objek), dan gerombol kelima 4,72% (9 objek). Alokasi objek setiap gerombol dapat dilihat pada Tabel 3.10.

Tabel 3.10: Alokasi Objek Pada Setiap Gerombol Untuk Model dengan Enam *Direct Effect*

Gerombol	1	2	3	4	5	Total
1	64,9263	0,7407	0,1580	0,0000	0,0000	68,8520
2	1,1895	62,2574	0,0000	0,0003	0,0000	63,4471
3	1,8604	0,0000	34,8420	0,0000	0,0000	36,7024
4	0,0236	0,0013	0,0000	18,9997	0,0000	19,0247
5	0,0003	0,0006	0,0000	0,0000	9,0000	9,0009
Total	68,0000	63,0000	35,0000	19,0000	9,0000	194,0000

### 3.6.4 Taksiran Parameter Model Terpilih

Pada Tabel 3.9 telah ditunjukkan nilai BIC dan *log-likelihood* masing-masing model. Penggerombolan terbaik adalah 5 gerombol dengan penambahan 6 *direct effect*. Setiap gerombol memiliki informasi yang berkaitan dengan peubah yang diamati pada setiap spesifikasi *smartphone* yang diamati. Nilai taksiran parameter setiap peubah dikelompokkan berdasarkan tipe data yang dijelaskan sebagai berikut:

1. Peubah Biner ( $\hat{\pi}_{ij}$ )

Untuk peubah biner dalam studi kasus ini terdapat pada peubah kemampuan mengakses dua sim sekaligus ( $X_2$ ), kemampuan menambah memori eksternal sebagai memori tambahan ( $X_7$ ), dan ada tidaknya *flash* ( $X_{11}$ ).

Parameter yang ditaksir dalam tipe peubah ini adalah  $\pi_{ij}$  yang merupakan nilai peluang ketika objek memberikan respon ya untuk peubah teramati ke- $i$  pada gerombol ke- $j$ . Hasil taksiran  $\pi_{ij}$  untuk 5 gerombol produk pada peubah  $X_2$ ,  $X_7$ , dan  $X_{11}$  ditunjukkan dalam Tabel 3.11.

Tabel 3.11: Taksiran Peluang Masing-Masing Kategori Pada Peubah Biner

Peubah	Resp.	(*)	1	2	3	4	5
Dual SIM $X_2$	Ya	$\hat{\pi}_{2j}$	<b>0,5694</b>	0,2023	<b>0,7254</b>	0,0562	0,0087
	Tidak	$1 - \hat{\pi}_{2j}$	0,4306	<b>0,7977</b>	0,2746	<b>0,9438</b>	<b>0,9913</b>
Mem. Eksternal $X_7$	Ya	$\hat{\pi}_{7j}$	<b>0,8831</b>	<b>0,7588</b>	<b>0,9991</b>	<b>0,9462</b>	0,0181
	Tidak	$1 - \hat{\pi}_{7j}$	0,1169	0,2412	0,0009	0,05338	<b>0,9819</b>
<i>Flash</i> $X_{11}$	Ya	$\hat{\pi}_{11j}$	<b>0,9989</b>	<b>0,9992</b>	0,3219	0,0612	<b>0,5600</b>
	Tidak	$1 - \hat{\pi}_{11j}$	0,0011	0,0008	<b>0,6781</b>	<b>0,9388</b>	0,4400

Keterangan :

Resp. : Respon kategori pada setiap peubah biner,

(\*) : Taksiran peluang masing-masing kategori pada setiap peubah biner.

Pada Tabel 3.11 dapat dilihat bahwa produk-produk yang dapat mengakses fitur dual SIM paling banyak tersebar di gerombol pertama dan ketiga. Pada gerombol keempat dan kelima hampir seluruhnya tidak dapat mengakses fitur tersebut, sedangkan pada gerombol kedua produk yang tidak dapat mengakses dual SIM lebih banyak dibandingkan dengan produk yang dapat mengakses dual SIM.

Gerombol pertama hingga keempat hampir seluruhnya dapat menggunakan memori eksternal sebagai perangkat memori tambahan. Produk pada gerombol kelima tidak memiliki tambahan *slot* memori eksternal. Hampir seluruh produk dalam gerombol pertama dan kedua dapat menggunakan *flash* pada kamera utamanya, pada gerombol keempat hampir seluruh produknya tidak terdapat *flash*. Pada gerombol ketiga dan gerombol kelima, produk yang terdapat *flash* lebih banyak daripada yang tidak terdapat *flash*.

2. Peubah Nominal ( $\pi_{ij(s)}$ )

Peubah nominal dalam studi kasus ini terletak pada peubah jenis SIM yang digunakan ( $X_3$ ), dan sistem operasi yang digunakan ( $X_5$ ). Parameter yang ditaksir dalam peubah ini adalah  $\pi_{ij(s)}$  yang merupakan nilai peluang ketika objek memberikan respon kategori ke- $s$  untuk peubah teramati ke- $i$  pada gerombol ke- $j$ . Hasil taksiran  $\pi_{ij(s)}$  untuk 5 gerombol produk untuk peubah  $X_3$  dan  $X_5$  ditunjukkan pada Tabel 3.12.

Tabel 3.12: Taksiran Peluang Masing-masing Kategori Pada Peubah Nominal

Peubah	Resp.	(*)	1	2	3	4	5
Jenis SIM $X_3$	Mini SIM	$\pi_{3j(1)}$	0,4033	0,0166	<b>0,6093</b>	0,2630	0,0061
	Micro SIM	$\pi_{3j(2)}$	<b>0,5964</b>	<b>0,7945</b>	0,3901	<b>0,7359</b>	0,0133
	Nano SIM	$\pi_{3j(3)}$	0,0003	0,1889	0,0006	0,0011	<b>0,9806</b>
Sistem Operasi $X_5$	Android	$\pi_{5j(1)}$	<b>0,7783</b>	<b>0,8844</b>	<b>0,8348</b>	<b>0,9447</b>	0,0175
	Windows	$\pi_{5j(2)}$	0,1760	0,0526	0,1647	0,0544	0,0025
	iOS	$\pi_{5j(3)}$	0,0305	0,0002	0,0003	0,006	<b>0,9794</b>
	BlackBerry	$\pi_{5j(4)}$	0,0153	0,0628	0,0001	0,0003	0,0006

Keterangan :

Resp. : Respon kategori pada setiap peubah nominal,

(\*) : Taksiran peluang masing-masing kategori pada setiap peubah nominal.

Pada Tabel 3.12 dapat dilihat bahwa pada gerombol pertama, kedua, dan keempat produk-produknya dominan menggunakan jenis Micro SIM. Untuk gerombol ketiga kebanyakan diisi oleh produk-produk yang menggunakan Mini SIM, dan gerombol kelima hampir seluruhnya hanya dapat menggunakan SIM bertipe Nano. Hampir di semua gerombol, dari gerombol pertama hingga keempat dipenuhi produk-produk bersistem operasi Android, sedangkan untuk gerombol kelima, semua produknya diisi produk-produk bersistem operasi iOS.

### 3. Peubah Ordinal ( $\hat{\pi}_{ij(s)} = \hat{\gamma}_{ij(s)} - \hat{\gamma}_{ij(s-1)}$ )

Peubah ordinal dalam studi kasus ini merupakan peubah terbanyak yang digunakan yakni 8 peubah antara lain pada peubah jaringan yang digunakan ( $X_1$ ), ukuran layar ( $X_4$ ), banyaknya *core* yang digunakan ( $X_6$ ), memori internal yang tersedia ( $X_8$ ), RAM yang tersedia ( $X_9$ ), besarnya kamera utama ( $X_{10}$ ), besarnya kamera depan ( $X_{12}$ ), dan kapasitas baterai ( $X_{15}$ ).

Parameter yang ditaksir dalam peubah ini adalah  $\gamma_{ij(s)} - \gamma_{ij(s-1)}$  yang merupakan nilai peluang ketika objek memberikan respon kategori ke- $s$  untuk peubah indikator ke- $i$  pada gerombol ke- $j$  jika diperhatikan urutan dari setiap tingkatan kategori. Hasil taksiran  $\gamma_{ij(s)} - \gamma_{ij(s-1)}$  untuk 5 gerombol produk pada peubah-peubah ordinal tersebut ditunjukkan dalam Tabel 3.13.

Pada peubah jaringan yang digunakan ( $X_1$ ), gerombol pertama, ketiga, dan keempat dominan menggunakan jaringan HSPA, sedangkan untuk gerombol kedua dan kelima hampir seluruh produknya menggunakan jaringan LTE. Untuk produk dengan spesifikasi jaringan EDGE hanya berada di gerombol ketiga.

Ukuran layar pada produk gerombol pertama berkisar antara 4,00 hingga 5,99 inch. Gerombol kedua didominasi oleh produk berukuran layar 5,00 hingga 5,99 inch. Pada gerombol ketiga berisikan produk berlayar kecil yakni kurang dari 4,00 inch. Gerombol keempat merupakan kumpulan produk berlayar besar yakni di atas 7,00 inch, serta pada gerombol kelima didominasi produk berukuran layar 5,00 hingga 5,99 inch.

Tabel 3.13: Taksiran Peluang Masing-masing Kategori Pada Peubah Ordinal

Peubah	Resp.	1	2	3	4	5
Jaringan $X_1$	EDGE	0,0050	0,0000	0,0989	0,0020	0,0000
	HSPA	<b>0,7259</b>	0,0773	<b>0,8810</b>	<b>0,5738</b>	0,0116
	LTE	0,2691	<b>0,9227</b>	0,0201	0,4243	<b>0,9884</b>
Ukuran Layar $X_4$	< 4,00 inch	0,0552	0,0143	0,4203	0,0000	0,0043
	4,00 – 4,99 inch	<b>0,5055</b>	0,2715	<b>0,5608</b>	0,0000	0,1468
	5,00 – 5,99 inch	0,4022	<b>0,5292</b>	0,0189	0,0017	<b>0,4746</b>
	6,00 – 6,99 inch	0,0331	0,1092	0,0000	0,0214	0,1639
	$\geq 7,00$ inch	0,0089	0,0759	0,0000	<b>0,9770</b>	0,2104
Jenis Core $X_6$	<i>Single Core</i>	0,0731	0,0080	0,3357	0,0068	0,1352
	<i>Dual Core</i>	<b>0,4626</b>	0,1506	<b>0,5839</b>	0,3156	<b>0,6128</b>
	<i>Quad Core</i>	0,4576	<b>0,7571</b>	0,0843	<b>0,6673</b>	0,2500
	<i>Octa Core</i>	0,0068	0,0843	0,0001	0,0102	0,0019
Mem. Internal $X_8$	4 GB	0,2505	0,0078	<b>0,7820</b>	0,0180	0,0000
	8 GB	<b>0,5100</b>	0,1328	0,2099	0,2291	0,0000
	16 GB	0,1809	0,3313	0,0079	<b>0,4091</b>	0,0023
	32 GB	0,0568	<b>0,4563</b>	0,0001	0,3245	0,1272
	64 GB	0,0017	0,0675	0,0000	0,0187	<b>0,4646</b>
	128 GB	0,0000	0,0043	0,0000	0,0006	0,4058
RAM $X_9$	512 MB	0,1261	0,0000	<b>0,7711</b>	0,0054	0,0331
	768 MB	0,0603	0,0002	0,0694	0,0104	0,0307
	1 GB	<b>0,7224</b>	0,0425	0,1567	<b>0,5005</b>	<b>0,7135</b>
	1,5 GB	0,0653	0,0790	0,0027	0,1813	0,1250
	2 GB	0,0255	<b>0,6324</b>	0,0002	0,2831	0,0945
	3 GB	0,0004	0,2146	0,0000	0,0187	0,0030
	4 GB	0,0000	0,0314	0,0000	0,0005	0,0000
Kamera Utama $X_{10}$	< 5,00 MP	0,0025	0,0000	<b>0,5681</b>	0,4159	0,0021
	5,00 – 7,99 MP	0,4131	0,0091	0,4292	<b>0,5776</b>	0,3809
	8,00 – 12,99 MP	<b>0,5411</b>	0,3168	0,0026	0,0065	<b>0,5657</b>
	$\geq 13,00$ MP	0,0443	<b>0,6741</b>	0,0000	0,0000	0,0513
Kamera Depan $X_{12}$	Tidak ada	0,0230	0,0003	<b>0,6467</b>	0,0132	0,0517
	0,01 – 1,99 MP	<b>0,7199</b>	0,2010	0,3523	<b>0,7795</b>	<b>0,8898</b>
	2,00 – 4,99 MP	0,2324	<b>0,4665</b>	0,0010	0,1897	0,0575
	$\geq 5,00$ MP	0,0247	0,3321	0,0000	0,0176	0,0010
Kapasitas Baterai $X_{15}$	$\leq 2000$ mAh	0,4435	0,1458	<b>0,9949</b>	0,0003	0,1202
	2001 – 3000 mAh	<b>0,4884</b>	<b>0,5488</b>	0,0051	0,0338	<b>0,4737</b>
	3001 – 4000 mAh	0,0620	0,2288	0,0000	0,4744	0,2587
	> 4000 mAh	0,0061	0,0766	0,0000	<b>0,4915</b>	0,1437

Keterangan :

Resp. : Respon kategori pada setiap peubah ordinal.

Produk-produk pada gerombol pertama terdiri dari produk dengan *processor Dual Core* dan *Quad Core*. Untuk gerombol ketiga dan kelima produk-produknya kebanyakan menggunakan *processor Dual Core*. Untuk gerombol kedua dan keempat didominasi oleh produk-produk yang menggunakan *processor Quad Core* dalam produknya.

Pada gerombol pertama dan keempat, memori *built-in* baik berupa memori internal maupun RAM kebanyakan menggunakan memori internal sebesar 8 GB dan RAM berkapasitas 1 GB. Untuk gerombol kedua didominasi oleh produk yang berkapasitas memori besar yakni 32 GB, dan memiliki RAM sebesar 2 GB. Pada gerombol kelima semua produknya menggunakan memori berkapasitas besar yakni 64 GB dan 128 GB, meskipun masih menggunakan RAM 1 GB. Gerombol ketiga mayoritas berisikan produk berkapasitas kecil yakni 4 GB untuk internal, dan 512 MB untuk RAM.

Untuk peubah kamera utama dan kamera depan, pada gerombol pertama dan keempat mayoritas produknya menggunakan kamera utama sebesar 8,00 hingga 12,99 MP, dan kamera depan sebesar VGA hingga 1,99 MP. Produk-produk dengan resolusi kamera yang tinggi terdapat di gerombol kedua yakni di atas 13 MP untuk kamera utama, dan di atas 2 MP untuk kamera depan. Lain halnya dengan gerombol ketiga, produk-produk di gerombol ini kebanyakan tidak memiliki kamera depan, dan kamera utamanya hanya sekitar dibawah 5,00 MP. Di gerombol kelima paling banyak memiliki kamera utama sebesar 8,00 hingga 12,99 MP dan kebanyakan hanya berkamera depan sebesar VGA hingga 1,99 MP.

Produk-produk pada gerombol keempat mengungguli gerombol yang lain dari sisi kapasitas baterai yakni di atas 3001 mAh. Produk-produk pada gerombol pertama, kedua, dan kelima memiliki kapasitas baterai tergolong menengah yakni 2001 mAh hingga 3000 mAh. Gerombol ketiga merupakan kumpulan produk dengan kapasitas baterai yang kecil, yakni hampir semuanya dibawah 2001 mAh.

#### 4. Peubah Kontinu Berdistribusi Normal

Pada studi kasus ini, peubah kontinu yang digunakan adalah peubah mengenai harga di pasaran pada bulan Mei 2015 , dan tingkat popularitas yang telah ditransformasi. Taksiran *mean* dan *varians* untuk peubah kontinu berdistribusi normal ini dapat dilihat pada Tabel 3.14.

Tabel 3.14: Taksiran *Mean* dan *Varians* Pada Peubah Kontinu

Peubah	(*)	1	2	3	4	5
ln Harga (ln $X_{13}$ )	$\mu_{13j}$	0,8040	1,5238	0,1014	0,9902	2,2488
	$e^{\mu_{13j}}$	<b>2,2344</b>	<b>4,5896</b>	<b>1,1067</b>	<b>2,6917</b>	<b>9,4763</b>
	$\sigma_{13}^2$	0,4559				
ln Pop. (ln $X_{14}$ )	$\mu_{14j}$	1,6246	2,0237	0,7678	0,6482	2,1707
	$e^{\mu_{14j}}$	<b>5,0763</b>	<b>7,5662</b>	<b>2,1550</b>	<b>1,9121</b>	<b>8,7644</b>
	$\sigma_{14}^2$	1,2761				

Keterangan:

(\*) : Taksiran *mean* dan *varians* untuk peubah kontinu berdistribusi normal.

Dapat dilihat pada Tabel 3.14, gerombol pertama didominasi oleh produk-produk dengan harga menengah, yakni rata-rata 2,2 juta rupiah. Gerombol kedua berisikan produk-produk dengan harga cukup mahal yakni 4,6 juta rupiah. Gerombol ketiga berisikan produk-produk yang relatif murah yakni hanya sekitar 1 juta rupiah. Pada gerombol keempat produk-produknya sekitar 2,7 juta rupiah, dan produk termahal terletak di gerombol

kelima dengan harga rata-rata 9,5 juta rupiah. Peubah harga ini memiliki nilai simpangan baku sebesar 0,6752.

Pada peubah tingkat popularitas ( $X_{14}$ ), gerombol kedua berisikan produk-produk yang cukup populer di mata masyarakat yakni sekitar 7,56%, gerombol pertama memiliki tingkat popularitas 5,08%. Tingkat popularitas untuk gerombol ketiga dan keempat masing-masing hanya sekitar 2%. Produk-produk yang digemari masyarakat paling banyak di gerombol kelima sebesar 8,76%. Simpangan baku untuk peubah tingkat popularitas ini sebesar 1,1296.

### 3.6.5 Pengklasifikasian Objek Pada Setiap Gerombol

Berdasarkan hasil penjelasan subbab sebelumnya, dengan metode *LCCA* didapatkan model 5 gerombol dengan penambahan 6 *direct effect*. Pemilihan model ini didasarkan pada kriteria BIC terkecil dan terpenuhinya asumsi kebebasan lokal.

Pengklasifikasian objek ke dalam gerombol didasarkan pada nilai taksiran parameter yang didapatkan menggunakan algoritma EM. Setiap gerombol yang dibentuk memiliki karakteristik yang berbeda-beda dari setiap peubah.

#### Gerombol Pertama

Pada gerombol pertama terdapat 68 buah handphone, dengan 18 produk (26,47%) diantaranya dari merek Samsung, 16 produk (19,12%) dari merek Lenovo, 11 produk (16,17%) dari merek Nokia, 10 produk (14,70%) dari merek Sony, 9 produk (13,23%) dari merek LG, 3 produk (4,41%) dari merek ASUS, 2 produk (2,94%) dari merek Apple, dan 1 produk (1,47%) dari merek Microsoft. Adapun

produk-produk yang termasuk dalam gerombol ini ditunjukkan oleh Tabel 3.15.

Pada gerombol ini sebanyak 18 produk (26,47%) sudah menggunakan jaringan LTE, dan 50 produk (73,53%) masih memakai HSPA. Produk dalam gerombol ini 44,11% (30 produk) tidak dapat menggunakan dual SIM. Jenis SIM yang digunakan didominasi oleh Micro SIM yakni sebanyak 40 produk (58,82%), Mini SIM sebanyak 28 produk (41,18%). Ukuran layar pada gerombol ini mayoritas berukuran 4,00-4,99 inch yaitu 37 produk (54,41%), 5,00-5,99 inch sebanyak 25 buah (36,76%), kurang dari 4 inch serta 6,00-6,99 inch ada 3 buah (4,41%).

Sistem operasi yang digunakan terbanyak oleh Android yaitu 53 buah (77,94%), Windows sebanyak 12 produk (17,65%), iOS sebanyak 2 buah (2,94%), dan BlackBerry sebanyak 1 buah (1,47%). Untuk *processor* yang digunakan sudah menggunakan *Dual Core* sebanyak 35 buah (51,47%), *Quad Core* sebanyak 28 buah (41,18%), *Single Core* sebanyak 4 buah (5,88%), dan *Octa Core* sebanyak 1 buah (1,47%) .

Pada gerombol ini terdapat 60 buah (88,23%) yang dapat menggunakan memori eksternal. Memori internal yang digunakan sebanyak 35 buah (51,47%) menggunakan 8GB, 19 buah (27,94%) menggunakan 4GB, 8 buah (11,76%) menggunakan 16GB, 5 buah (7,35%) menggunakan 32GB, dan hanya 1 buah yang menggunakan 64 GB (1,47%). RAM yang digunakan pada gerombol ini, 48 produk (70,59%) menggunakan RAM 1GB, 10 produk (14,70%) dengan RAM 512MB, 5 produk (7,35%) dengan RAM 1,5GB, 4 produk (5,88%) dengan RAM 768GB, dan RAM 2GB sebanyak 1 produk (1,47%).

Tabel 3.15: Alokasi Produk *Smartphone* Pada Gerombol Pertama

No	Merek	Nama Produk	No	Merek	Nama Produk
1	Apple	iPhone 4	35	Nokia	XL DS
2	Apple	iPhone 4s	36	Nokia	Lumia 925
3	ASUS	Zenfone C	37	Nokia	Lumia 920
4	ASUS	Zenfone 4s	38	Nokia	Lumia 625
5	ASUS	Zenfone 5	39	Nokia	Lumia 820
6	BlackBerry	Z3	40	Nokia	Lumia 1320
7	Lenovo	A706	41	Samsung	Galaxy Ace 4
8	Lenovo	S820	42	Samsung	Galaxy Ace 3
9	Lenovo	S890	43	Samsung	Galaxy S3 Mini
10	Lenovo	P780	44	Samsung	Galaxy S2
11	Lenovo	A850	45	Samsung	Galaxy S3
12	Lenovo	S920	46	Samsung	Galaxy Ativ S
13	Lenovo	A536	47	Samsung	Galaxy S duos 2
14	Lenovo	A859	48	Samsung	Galaxy Core
15	Lenovo	S930	49	Samsung	Galaxy Grand
16	Lenovo	S939	50	Samsung	Galaxy Grand Neo Plus
17	Lenovo	S860	51	Samsung	Grand Neo
18	Lenovo	S856	52	Samsung	Galaxy J1
19	Lenovo	A6000	53	Samsung	Galaxy Core II
20	LG	Optimus L7	54	Samsung	Galaxy Mega 5,8
21	LG	Optimus L9 II	55	Samsung	Galaxy S4 Mini
22	LG	Optimus L7 II	56	Samsung	Galaxy Ativ I930
23	LG	L80 Dual	57	Samsung	Galaxy Grand Prime
24	LG	L Fino	58	Samsung	Grand 2
25	LG	L Bello	59	Sony	Xperia J
26	LG	G3 Stylus	60	Sony	Xperia L
27	LG	G Pro Lite	61	Sony	Xperia P
28	LG	G2 Mini	62	Sony	Xperia M
29	Microsoft	Lumia 535 DS	63	Sony	Xperia C
30	Nokia	Lumia 620	64	Sony	Xperia V
31	Nokia	Lumia 900	65	Sony	Xperia E3
32	Nokia	Lumia 720	66	Sony	Xperia M2
33	Nokia	X2 DS	67	Sony	Xperia M2 Aqua
34	Nokia	Lumia 730 DS	68	Sony	Xperia T3

Kamera utama dengan resolusi 8,00 hingga 12,99 MP sebanyak 37 buah (54,41%) pada gerombol ini semuanya dilengkapi *flash*, untuk kamera kisaran 5 MP hingga 7,99 MP sebanyak 29 buah (42,64%), dan di atas 13 MP sebanyak 2 buah (2,94%). Kamera depan yang digunakan memakai resolusi tinggi di atas 5MP hanya sebanyak 3 buah (4,41%), 2MP hingga 4,99MP sebanyak 10 buah (14,70%), dan VGA hingga 1,99MP sebanyak 55 buah (80,88%).

Produk-produk dalam gerombol ini terdiri dari 32 produk (47,05%) dengan baterai berkapasitas  $\leq 2000$  mAh, 31 produk (45,58%) berkapasitas 2001 hingga 3000 mAh, dan 5 buah (7,35%) berkapasitas 3001 hingga 4000 mAh. Harga dalam gerombol ini memegang harga yang menengah di pasaran yakni sekitar 2,2 juta rupiah, dan memiliki tingkat popularitas sebesar 5,05%.

Dari berbagai penjelasan di atas, dapat disimpulkan bahwa pada gerombol pertama yang terdiri dari 68 buah produk *smartphone* ini, merupakan gerombol yang berisikan produk-produk dengan spesifikasi menengah atau disebut produk *medium* yang rata-rata dimiliki oleh masyarakat.

## **Gerombol Kedua**

Dari 194 produk yang dianalisis, sebanyak 63 produk termasuk dalam gerombol kedua. Dari 63 produk tersebut, 60 diantaranya dalam bentuk *handphone* dan tiga produk dalam bentuk *tablet*. Adapun yang termasuk dalam gerombol ini sebanyak 19 produk (30,16%) dari merek Samsung , 12 produk (19,05%) dari merek Sony , 12 (19,05%) produk dari merek Lenovo , 8 produk (12,70%) dari merek LG , 5 produk (7,93%) dari merek ASUS, 4 produk (6,35%) dari merek BlackBerry , dan 3 produk (4,76%) dari merek Nokia . Adapun produk-produk yang termasuk dalam gerombol ini ditunjukkan oleh Tabel 3.16.

Tabel 3.16: Alokasi Produk *Smartphone* Pada Gerombol Kedua

No	Merek	Nama Produk	No	Merek	Nama Produk
1	ASUS	Zenfone 6	33	Samsung	Galaxy A3
2	ASUS	Zenfone Zoom	34	Samsung	Galaxy E5 Duos
3	ASUS	Zenfone 2	35	Samsung	Galaxy Mega 2
4	ASUS	Padfone S	36	Samsung	Galaxy S4 Zoom
5	ASUS	Zenfone 2s	37	Samsung	Galaxy Mega 6,3
6	BlackBerry	Q5	38	Samsung	Galaxy E7
7	BlackBerry	Z10	39	Samsung	Galaxy S4 Active
8	BlackBerry	Q10	40	Samsung	Galaxy A5
9	BlackBerry	Z30	41	Samsung	Galaxy Note 3 Neo
10	Lenovo	K3 Note	42	Samsung	Galaxy K Zoom
11	Lenovo	Golden Warrior	43	Samsung	Galaxy S4
12	Lenovo	A7000	44	Samsung	Galaxy Tab S 8,4
13	Lenovo	Vibe X	45	Samsung	Galaxy Note 2
14	Lenovo	K900	46	Samsung	Galaxy S5
15	Lenovo	P70	47	Samsung	Galaxy Tab S 10,5
16	Lenovo	Vibe X2	48	Samsung	Galaxy Note 3
17	Lenovo	Vibe Z	49	Samsung	Galaxy Alpha
18	Lenovo	Livo S90	50	Samsung	Galaxy Note 4
19	Lenovo	P90	51	Samsung	Galaxy S6 Edge
20	Lenovo	Vibe Shot	52	Sony	Xperia T2 Ultra
21	Lenovo	Vibe Z2 Pro	53	Sony	Xperia C3
22	LG	Nexus 5	54	Sony	Xperia ZR
23	LG	Optimus G	55	Sony	Xperia Z
24	LG	G2	56	Sony	Xperia Z Ultra
25	LG	Optimus G Pro	57	Sony	Xperia ZL
26	LG	G Pro 2	58	Sony	Xperia Z1 Compact
27	LG	G3	59	Sony	Xperia Z1
28	LG	G Flex	60	Sony	Xperia Z3 Compact
29	LG	G Flex 2	61	Sony	Xperia Z2A
30	Nokia	Lumia 1020	62	Sony	Xperia Z3 Tablet
31	Nokia	Lumia 930	63	Sony	Xperia Z3
32	Nokia	Lumia 1520			

Pada gerombol ini sebanyak 58 produk (92,06%) sudah menggunakan jaringan LTE, dan hanya 5 produk (7,93%) yang masih memakai HSPA. Produk dalam gerombol ini 79,36% (50 produk) tidak dapat menggunakan dual SIM, dan 13 produk (20,63%) menggunakan dual SIM. Jenis SIM yang digunakan didominasi oleh Micro SIM yakni sebanyak 50 produk (79,36%), lalu Nano SIM sebanyak 12 produk (19,05)%, dan Mini SIM hanya 1 produk (1,59%). Ukuran layar pada gerombol ini mayoritas berukuran 5,00-5,99 inch yaitu 39 produk (61,90%), 4,00-4,99 inch sebanyak 11 buah (17,46%), 6,00-6,99 inch sebanyak 8 produk (12,70%), lebih dari 7 inch sebanyak 3 produk (4,76%), serta kurang dari 4 inch ada 2 buah (3,17%).

Sistem operasi yang terbanyak digunakan pada produk-produk di gerombol ketiga adalah Android. Untuk *processor* yang digunakan sudah menggunakan *Quad Core* sebanyak 50 buah (79,37%), *Dual Core* sebanyak 9 buah (14,28%), dan *Octa Core* sebanyak 4 buah (6,34%) .

Pada gerombol ini terdapat 48 buah (76,19%) yang dapat menggunakan memori eksternal. Memori internal yang digunakan sebanyak 28 buah (44,44%) menggunakan 32GB, 24 buah (38,09%) menggunakan 16GB, 7 buah (11,11%) menggunakan 8GB, dan hanya 4 buah yang menggunakan 64 GB (6,34%).

RAM yang digunakan pun pada gerombol ini yaitu 40 produk (63,49%) menggunakan RAM 2GB, 14 produk (22,22%) dengan RAM 3GB, 5 produk (7,94%) dengan RAM 1,5GB dan RAM 1GB, serta RAM 4GB sebanyak 2 produk (3,17%).

Kamera utama dengan resolusi tinggi di atas 13 MP sebanyak 44 buah (69,84%) pada gerombol ini yang semuanya dilengkapi *flash*, untuk kamera kisaran 8 MP hingga 12,99 MP sebanyak 18 buah (28,57%), dan 5 MP hingga 7,99 MP sebanyak 1 buah (1,58%). Kamera depan yang digunakan memakai resolusi tinggi

di atas 5MP sebanyak 20 buah (31,74%), 2MP hingga 4,99MP sebanyak 32 buah (50,79%), dan VGA hingga 1,99MP sebanyak 11 buah (17,46%).

Produk-produk dalam gerombol ini terdiri dari 40 produk (63,49%) dengan baterai berkapasitas 2001 hingga 3000 mAh, 15 produk (23,80%) berkapasitas 3001 hingga 4000 mAh, 5 buah (7,94%) berkapasitas dibawah 2001 mAh, dan 3 buah (4,76%) di atas 4000 mAh. Harga dalam gerombol ini memegang harga tinggi di pasaran yakni sekitar 4,6 juta rupiah, dan memiliki tingkat popularitas sebesar 7,56%.

Dari berbagai penjelasan di atas, dapat disimpulkan bahwa pada gerombol kedua yang terdiri dari 63 buah produk *smartphone* ini, merupakan gerombol yang berisikan produk dengan spesifikasi tinggi atau disebut juga produk *premium/high-end* yang rata-rata diminati masyarakat.

### **Gerombol Ketiga**

Pada gerombol ketiga, terdapat 35 buah produk yang semuanya berbentuk *handphone*. Dari 35 produk tersebut diantaranya 9 produk (25,71%) dari merek Samsung, 7 produk (20%) dari merek LG, 6 produk (17,14%) dari merek Nokia, 6 produk (17,14%) dari merek Lenovo, 4 produk (11,43%) dari merek Sony, dan 1 produk (2,86%) dari merek ASUS. Adapun produk-produk yang termasuk dalam gerombol ini ditunjukkan oleh Tabel 3.17.

Pada gerombol ini sebanyak 32 produk (91,43%) menggunakan jaringan HSPA, dan masih terdapat 3 produk (8,57%) yang menggunakan jaringan EDGE. Produk dalam gerombol ini 74,28% (26 produk) dapat menggunakan dual SIM, dan 9 produk (25,71%) yang tidak menggunakan dual SIM. Jenis SIM yang digunakan didominasi oleh Mini SIM yakni sebanyak 21 produk (60%), dan Micro SIM sebanyak 14 produk (40%). Ukuran layar pada gerombol ini mayoritas berukuran

4,00-4,99 inch yaitu 20 produk (57,14%), serta 15 buah (42,86%) yang memiliki spesifikasi layar kurang dari 4 inch.

Tabel 3.17: Alokasi Produk *Smartphone* Pada Gerombol Ketiga

No	Merek	Nama Produk	No	Merek	Nama Produk
1	ASUS	Zenfone 4	18	Nokia	X+
2	Lenovo	A316	19	Nokia	Lumia 520
3	Lenovo	A390	20	Nokia	Lumia 530 Duos
4	Lenovo	A369	21	Nokia	Lumia 525
5	Lenovo	Muszik A319	22	Nokia	Lumia 630 Duos
6	Lenovo	A516	23	Samsung	Galaxy Young 2
7	Lenovo	A526	24	Samsung	Galaxy Star
8	LG	Optimus L1 2	25	Samsung	Galaxy Chat
9	LG	A690	26	Samsung	Galaxy Star Deluxe
10	LG	L40 Dual	27	Samsung	Galaxy Pocket Neo
11	LG	Optimus L3 II	28	Samsung	Galaxy young S6310
12	LG	L60 Dual	29	Samsung	Galaxy Star Pro Duos
13	LG	Optimus L4	30	Samsung	Galaxy V
14	LG	Optimus L5	31	Samsung	Galaxy Fame
15	Microsoft	Lumia 435 Duos	32	Sony	Xperia E
16	Microsoft	Lumia 532 Duos	33	Sony	Xperia Miro
17	Nokia	X Duos	34	Sony	Xperia Tipo
			35	Sony	Xperia Go

Sistem operasi yang digunakan pun terbanyak oleh Android yaitu 29 buah (82,85%) dan Windows sebanyak 6 buah (17,14%). Untuk *processor* yang digunakan sudah menggunakan *Dual Core* sebanyak 16 buah (45,71%), *Single Core* sebanyak 15 buah (42,86%), dan *Quad Core* sebanyak 4 buah (11,42%) .

Pada gerombol ini semuanya dapat menggunakan memori eksternal. Memori internal yang digunakan sebanyak 27 buah (77,14%) hanya menggunakan 4GB, dan 8 buah yang menggunakan 8 GB (22,86%). RAM yang digunakan pun pada gerombol ini yaitu 27 produk (77,14%) hanya menggunakan RAM 512MB, 5 produk (14,28%) dengan RAM 1GB, dan 3 buah produk (8,57%) dengan RAM 768MB.

Kamera utama terbatas hanya dibawah 5 MP yakni sebanyak 21 produk (60%), dan 14 produk menggunakan kamera 5,00 hingga 7,99 MP. Produk pada gerombol ini yang dilengkapi *flash* hanya 10 produk (28,57%), dan 25 produk yang lain tidak menggunakan *flash* (71,42%). Kamera depan yang digunakan memakai resolusi rendah yakni VGA hingga 1,99 MP sebanyak 9 buah (25,71%), 2,00 hingga 4,99 MP sebanyak 1 buah (2,85%), dan paling banyak tidak terdapat kamera depan yakni 25 buah (71,42%).

Produk-produk dalam gerombol ini semuanya berkapasitas di bawah 2000 mAh. Di pasaran, harga untuk produk-produk pada gerombol ini tergolong rendah yakni sekitar 1,1 juta rupiah, dan memiliki tingkat popularitas hanya sebesar 2,15%.

Dari berbagai penjelasan di atas, dapat disimpulkan bahwa pada gerombol ketiga yang terdiri dari 35 buah produk *smartphone* ini, merupakan gerombol yang berisikan produk dengan spesifikasi rendah atau disebut juga produk *low-end* yang kurang diminati masyarakat.

### **Gerombol Keempat**

Pada gerombol keempat terdiri dari 19 buah produk *smartphone* dimana 11 produk (57,89%) dari merek Samsung, 3 produk (15,78%) dari merek ASUS, 3 produk (15,78%) dari merek Lenovo , 1 produk (5,26%) dari merek LG, dan 1 produk (5,26%) dari merek Nokia. Adapun produk-produk yang termasuk dalam gerombol ini ditunjukkan oleh Tabel 3.18.

Pada gerombol ini sebanyak 9 buah (47,37%) menggunakan jaringan HSPA, 9 buah (47,47%) sudah memakai LTE, dan 1 buah (5,26%) menggunakan EDGE. Produk dalam gerombol ini hampir semuanya tidak terdapat dual SIM yaitu 18 produk (94,74%), hanya 1 produk saja yang menggunakan dual

Tabel 3.18: Alokasi Produk *Smartphone* Pada Gerombol Keempat

No	Merek	Nama Produk	No	Merek	Nama Produk
1	ASUS	Fonepad 7	11	Samsung	Galaxy Tab 3 7 Lite
2	ASUS	Google Nexus	12	Samsung	Galaxy Tab 3 7,0
3	ASUS	Fonepad 8	13	Samsung	Galaxy Tab 3 8,0
4	Lenovo	A3300	14	Samsung	Galaxy Tab 4 7,0
5	Lenovo	A3500	15	Samsung	Galaxy Tab 4 8,0
6	Lenovo	A5500	16	Samsung	Galaxy Tab 3 8,0
7	LG	G Pad 8,0	17	Samsung	Galaxy Note 8
8	Nokia	Lumia 2520	18	Samsung	Galaxy Tab 4 10,1
9	Samsung	Galaxy Tab 2 7,0	19	Samsung	Galaxy Note 10,1
10	Samsung	Galaxy Tab 3V			

SIM (5,26%). Jenis SIM yang digunakan didominasi oleh Micro SIM yakni 14 buah (73,68%), dan Mini SIM 5 buah (26,31%). Ukuran layar pada gerombol ini semuanya berukuran di atas 7,00 inch. Sistem operasi yang digunakan pun terbanyak oleh Android yaitu 18 buah (94,74%), dan Windows 1 buah (5,26%).

Untuk *processor* yang digunakan sudah menggunakan *Quad Core* sebanyak 13 buah (68,42%), dan *Dual Core* sebanyak 6 buah (31,58%). Pada gerombol ini hampir semuanya dapat mengakses memori eksternal sebagai memori tambahan yaitu 18 produk (94,74%). Memori internal yang digunakan sebanyak 9 produk (47,37%) menggunakan 16GB, 6 produk (31,57%) menggunakan 32GB, dan 4 buah (21,06%) menggunakan 8GB. RAM yang digunakan pada produk-produk di gerombol ini sebanyak 9 produk (47,37%) menggunakan RAM 1GB. Sementara itu produk dengan spesifikasi RAM 1,5GB dan 2 GB masing-masing sebanyak 5 buah (26,32%).

Kamera utama dengan resolusi rendah dibawah 5MP sebanyak 8 produk (42,10%), dan resolusi 5,00-7,99 MP sebanyak 11 produk (57,90%). Kemampuan *flash* dalam gerombol ini hanya tersedia pada 1 buah produk saja (5,26%), dan tidak tersedia *flash* sebanyak 18 buah (94,74%). Kamera depan yang digunakan

pun hanya VGA hingga 1,99MP yakni sebanyak 13 produk (68,42%), 2MP hingga 4,99MP sebanyak 5 buah (26,31%), dan ada 1 produk yang tidak menggunakan kamera depan (5,26%).

Kapasitas baterai pada gerombol ini memiliki kapasitas yang cukup tinggi yaitu 10 produk (52,63%) menggunakan baterai berkapasitas 3001 hingga 4000 mAh, dan 9 produk (47,337%) menggunakan baterai di atas 4000 mAh. Harga dalam gerombol ini memegang harga yang menengah di pasaran yakni sekitar 2,7 juta rupiah, dan memiliki tingkat popularitas yang hanya sebesar 1,91%.

Dari berbagai penjelasan di atas, dapat disimpulkan bahwa pada gerombol keempat yang terdiri dari 19 buah produk *smartphone* ini, merupakan gerombol yang berisikan produk-produk dengan spesifikasi tablet berukuran besar, yang digunakan masyarakat untuk menunjang produktivitas atau hiburan.

### Gerombol Kelima

Pada gerombol kelima terdiri dari 9 buah produk. Semua produk dalam gerombol terakhir ini berisikan produk-produk Apple yang memiliki harga sangat tinggi. Adapun produk-produk yang termasuk dalam gerombol ini ditunjukkan oleh Tabel 3.19.

Tabel 3.19: Alokasi Produk *Smartphone* Pada Gerombol Kelima

No	Merek	Nama Produk	No	Merek	Nama Produk
1	Apple	iPhone 5	6	Apple	iPad Air
2	Apple	iPhone 5c	7	Apple	iPad Air 2
3	Apple	iPhone 5s	8	Apple	iPad Mini 2
4	Apple	iPhone 6	9	Apple	iPad Mini 3
5	Apple	iPhone 6 Plus			

Pada gerombol ini semua produknya sudah memakai LTE, serta tidak dapat menggunakan dual SIM, dan jenis SIM yang digunakan berbentuk Nano

SIM. Ukuran layar pada gerombol ini 4 buah (44,44%) berukuran 4,00 hingga 4,99 inch, 4 buah (44,44%) di atas 7 inch, dan 1 buah (11,11%) antara 5,00 hingga 5,99 inch. Sistem operasi yang digunakan semuanya menggunakan sistem operasi iOS.

Untuk *processor* yang digunakan hampir semuanya menggunakan *Dual Core* yaitu 8 produk (88,88%) dan hanya 1 yang menggunakan *Quad Core* (11,11%) . Pada gerombol ini semuanya tidak dapat mengakses memori eksternal sebagai memori tambahan. Memori internal yang digunakan sebanyak 4 buah (44,44%) menggunakan 64GB, 4 buah (44,44%) menggunakan kapasitas super besar yaitu 128GB, dan hanya 1 produk (11,11%) yang menggunakan 32GB. RAM yang digunakan pada gerombol ini hampir semuanya menggunakan RAM 1GB yaitu 88,88% dan 1 buah (11,11%) menggunakan RAM 2GB.

Kamera utama dengan resolusi 5,00-7,99 MP sebanyak 3 buah (33,33%), dan 8,00-12,99MP sebanyak 6 buah (66,66%), dimana 5 buah (55,55%) dilengkapi *flash*, dan 4 buah (44,44%) tidak dilengkapi *flash*. Kamera depan yang digunakan pun semuanya hanya VGA hingga 1,99MP.

Pada gerombol ini kapasitas baterai yang digunakan yaitu dibawah 2000 mAh sebanyak 4 buah (44,44%), di atas 4000 mAh sebanyak 4 buah (44,44%), dan 1 buah memakai 2001-3000 mAh (11,11%). Harga produk pada gerombol ini di pasaran tergolong tinggi yaitu sekitar 9.5 juta rupiah, dan memiliki tingkat popularitas yang tinggi sebesar 8,76%.

Dari berbagai penjelasan di atas, dapat disimpulkan bahwa pada gerombol kelima yang terdiri dari 9 buah produk *smartphone* dari produk Apple ini, merupakan gerombol yang berisikan produk-produk dengan spesifikasi menengah namun harganya sangat tinggi, yang biasanya menunjukkan tingkat sosial penggunaannya.

### 3.6.6 Rasio Keragaman Gerombol Model Terpilih

Setelah dilakukan pemilihan model dengan BIC terkecil dan terpenuhinya asumsi kebebasan lokal, maka dapat dihitung rasio keragaman dari gerombol model terpilih.

Untuk rasio keragaman pada peubah kontinu, yakni pada peubah harga ( $X_{13}$ ) dan tingkat popularitas ( $X_{14}$ ) menggunakan rumus yang terdapat pada subbab 2.13. Nilai SSW, SSB, dan rasio keragaman dari kedua peubah kontinu ini dapat dilihat pada Tabel 3.20.

Tabel 3.20: Rasio Keragaman Pada Peubah Kontinu

No	Peubah	SSW	SSB	Rasio Keragaman
1	Harga ( $X_{13}$ )	1,5283	11,2808	0,1354
2	Popularitas ( $X_{14}$ )	6,4685	23,7902	0,2719

Dapat dilihat dari Tabel 3.20 pada peubah harga ( $X_{13}$ ) memiliki nilai SSW dan SSB lebih rendah daripada peubah tingkat popularitas ( $X_{14}$ ). Dilihat dari nilai SSW peubah harga yang lebih rendah daripada peubah tingkat popularitas, untuk semua gerombol, harga objek-objek pada gerombol yang sama relatif homogen daripada tingkat popularitas yang cenderung agak beragam. Dilihat dari nilai SSB peubah harga yang lebih rendah daripada peubah tingkat popularitas, harga antar setiap gerombol lebih homogen daripada tingkat popularitas antar gerombol yang cenderung lebih beragam.

Nilai SSW, SSB, dan rasio keragaman untuk kedua belas peubah kategorik yang digunakan dapat dilihat pada Tabel 3.21.

Berdasarkan Tabel 3.21 dapat dilihat rasio keragaman tertinggi terdapat pada peubah kapasitas baterai ( $X_{15}$ ) yaitu sebesar 0,0997, dan rasio keragaman terendah pada peubah ketersediaan *slot* memori eksternal ( $X_7$ ) sebesar 0,022.

Tabel 3.21: Rasio Keragaman Pada Peubah Kategorik

No	Peubah	SSW	SSB	Rasio Keragaman
1	Jaringan ( $X_1$ )	0,3694	14,2778	0,0259
2	Dual SIM ( $X_2$ )	0,4286	12,9651	0,0331
3	Jenis SIM ( $X_3$ )	0,4546	12,7449	0,0356
4	Ukuran Layar ( $X_4$ )	0,5055	11,3702	0,0444
5	Sistem Operasi ( $X_5$ )	0,3606	14,1938	0,0254
6	Jenis <i>Core</i> ( $X_6$ )	0,4900	12,1120	0,0405
7	<i>Ext. Memory</i> ( $X_7$ )	0,3207	14,5927	0,0220
8	<i>Int. Memory</i> ( $X_8$ )	0,5477	10,5120	0,0521
9	Kapasitas RAM ( $X_9$ )	0,4962	11,9355	0,0416
10	Kamera Utama ( $X_{10}$ )	0,4950	12,0580	0,0411
11	<i>flash</i> ( $X_{11}$ )	0,2336	5,6982	0,0409
12	Kamera Depan ( $X_{12}$ )	0,4712	4,9827	0,0946
13	Kapasitas Baterai ( $X_{15}$ )	0,4790	4,9030	0,0977

Setelah dilakukan perhitungan untuk rasio keragaman pada masing-masing peubah kontinu dan kategorik, selanjutnya adalah menentukan nilai rasio keragaman total. Rasio keragaman total diperoleh dengan cara mencari rata-rata rasio keragaman dari seluruh peubah, baik peubah kontinu maupun kategorik. Setelah dihitung rata-ratanya maka diperoleh nilai rasio keragaman total dalam model terpilih ini adalah sebesar 0,0667.

Nilai rasio keragaman sebesar 0.0667 (kurang dari 1) ini menunjukkan bahwa model yang dipilih sudah baik dalam menggerombolkan 194 objek. Model terpilih tersebut adalah model dengan 5 gerombol dan 6 *direct effect*. Tingkat akurasi untuk model ini berdasarkan rumus pada subbab 2.13 didapatkan sebesar 93,33%.

# BAB IV

## PENUTUP

### 4.1 Kesimpulan

1. Model *Latent Class Cluster Analysis* pertama kali diperkenalkan oleh Lazarsfeld pada tahun 1963 yang terbatas pada peubah biner saja, lalu dikembangkan oleh Moustaki dan Papagerogiou (2004) untuk mengklasifikasikan objek atau reponden ke dalam kelas laten berdasarkan peubah bertipe campuran. Model *Latent Class Cluster Analysis* untuk peubah bertipe campuran adalah:

$$f(x_h) = \sum_{j=1}^J \eta_j \left[ \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}} (1 - \pi_{ij})^{1-x_{ih}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}} \right] \right. \\ \left. \left[ \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}} \right] \left[ \frac{1}{\prod_{i=1}^Q \sqrt{2\pi}\sigma_i} \exp \left[ -\sum_{i=1}^Q \frac{1}{2\sigma_i^2} (x_{ih} - \mu_{ij})^2 \right] \right] \right] \right]$$

2. Penaksiran parameter yang digunakan dalam model ini menggunakan metode maksimum *Likelihood* dengan algoritma EM (*Expectation Maximization*) hingga tercapai nilai taksiran parameter yang konvergen atau  $|\hat{\theta}_{(t)} - \hat{\theta}_{(t-1)}|$  mendekati nol.
3. Penggerombolan terhadap 194 produk ponsel pintar (*smartphone*) dengan menggunakan *Latent Class Cluster Ananlysis (LCCA)* menghasilkan 5 gerombol kelas laten. Model dengan 5 gerombol ini memiliki BIC terkecil yaitu

5041,62, memiliki nilai BVR (*Bivariate Residual*) untuk setiap pasangan peubah kategorik sebesar 0-6,42, memiliki rasio keragaman sebesar 0,0667, serta memiliki tingkat akurasi sebesar 93,33%.

4. Pada gerombol pertama berisikan 68 produk dengan spesifikasi menengah dengan harga yang menengah pula, gerombol kedua berisikan 63 produk berspesifikasi tinggi dengan harga yang cukup tinggi, gerombol ketiga berisikan 35 produk berspesifikasi rendah yang memiliki harga relatif murah, gerombol keempat berisikan 19 produk dengan spesifikasi tablet dengan harga-harga menengah kebawah, gerombol terakhir berisikan 9 produk *smart-phone* yang memiliki harga paling mahal, namun memiliki spesifikasi menengah.

## 4.2 Saran

Berdasarkan penelitian yang telah dilakukan, ada beberapa saran yang dapat disampaikan, yaitu:

- Bagi pembaca yang tertarik pada penelitian menggunakan *LCCA* ini, dapat melakukan penelitian mengenai metode struktur laten yang lain seperti model analisis faktor kelas laten (*Latent Class Factor Analysis/LCFA*) dan model analisis regresi kelas laten (*Latent Class Regression Analysis/LCRA*).
- Dalam pembentukan model *LCCA* dapat menggunakan model distribusi keluarga eksponensial yang lain, seperti Poisson, Beta, Gamma, dan sebagainya.
- Dalam penanganan asumsi kebebasan lokal, dapat diterapkan metode lain selain penambahan *direct effect*, seperti pereduksian peubah.

## DAFTAR PUSTAKA

- Agresti, Alan. 2002. *Categorical Data Analysis*. Edisi ke-2. New Jersey: Jason Willey & Sons Inc.
- Budiati, Shinta. 2014. Pengelompokan Daerah Penghasil Bahan Dasar Tepung Komposit di Indonesia Menggunakan *Latent Class Cluster Analysis*. Skripsi. Surakarta: Universitas Negeri Sebelas Maret.
- Celeux, G., Biernacki, C., & Govaert, G. 1997. *Choosing Models in Model Based Clustering and Discriminant Analysis*. Technical Report. Rhone Alpes: INRIA.
- Dempster, A. P. 1977. *Maximum Likelihood from Incomplete Data via EM Algorithm*. *Journal of the Royal Statistic Society. Series B (Methodical)*.
- Dewi, A., Sutikno dan Kuswanto, H. 2012. *Metode Cluster Ensembl* untuk Pengelompokan Desa Perdesaan di Provinsi Riau. Tugas Akhir. Surabaya: Institut Teknologi Sepuluh November (ITS).
- Everitt, B. S. 1993. *Cluster Analysis*. London: Edward Arnold.
- Fraley, C. 2002. *Model Based Clustering, Discriminant Analysis, and Density Estimation*. *Journal of the American Statistical Association*, Vol. 97, pp 611-631.
- Hagenaars, J. A dan McCuthceon, A. L. 2002. *Applied Latent Class Analysis*. UK: Cambridge University Press.

- Kaufman, L. dan Rouseeuw, P. J. 1990. *Finding Group in Data: An Introduction to Cluster Analysis*. New Jersey: Jason Willey & Sons Inc.
- Handayani, Novalita. 2011. Pendugaan Parameter *Latent Class Model* dengan Menggunakan Algoritma EM. Skripsi. Jakarta: Universitas Negeri Jakarta.
- Hidayat, Yayat. 2009. Perbandingan Teknik Klasifikasi *Latent Class Cluster* dan *K-Means*. *Tesis*. Bandung: Universitas Padjajaran.
- Lazarsfeld, P. F. dan Henry, N.W. 1968. *Latent Structure Analysis*. Boston: Houghton Mifflin.
- Mansur, Asep Arifin. 2009. Analisis *Latent Class Cluster* dalam Pengklasifikasian Subjek Penerima BLT . *Tesis*. Bandung: Universitas Padjajaran.
- Maulidya, F.D. 2009. Penerapan Analisis *Latent Class Cluster* (LCC) dalam Segmentasi Pasar. Skripsi. Malang: Universitas Brawijaya.
- Moustaki, Irini. 1996. *A Latent Trait and a Latent Class Model for Mixed Observed Variable*, *Journal of Mathematical and Statistical Psychology* Vol. 49, pp 313-334.
- Moustaki, I. dan Papageorgiou, I. 2005. *Latent Class Models for Mixed Outcomes with Application in Archaeometry, Computational Statistics and Data Analysis* Vol. 48(3), pp 659-3675.
- Nainggolan, B. M. H. 2009. Perbandingan Analisis Laten Kelas dengan Kriteria WHO untuk Penggerombolan Pasien Demam Dangu (DD) dan Demam Berdarah Dangu (DBD) . *Tesis*. Bogor: Institut Pertanian Bogor.
- Novianti. 2008. *Latent Class Model*. Skripsi. Depok: Universitas Indonesia.

- Quinn, K. 2001. *The Newton Raphson Algorithm for Function Optimization*. USA: University of Washington.
- Razali, N. M. dan Wah, Y.B. 2011. *Power Comparasion of Saphiro-Wilk, Kolmogorov-Smirnov, Liliefors, and Anderson Darling*. New York: Jason Willey & Sons Inc.
- Rendra, Gusti. 2014. Penerapan *Latent Class Cluster Analysis* Pada Data Responden Pengguna *Flexy Trendy* di Kabupaten Jember. Skripsi. Jember: Universitas Jember.
- Royston, J. P. 1992. *Approximating The Saphiro-Wilk for Non-normality*. New York: Jason Willey & Sons Inc.
- Sambandam. 2010. *Cluster Analysis Gets Complicated*. Berlin: Marketing Research.
- Sartono, B dkk. 2003. Analisis Peubah Ganda. Bogor: Institut Pertanian Bogor.
- Sharma, Shubhash. 1996. *Applied Multivariate Techniques*. New York: Jason Willey & Sons Inc.
- Sukestiyarno dan Budi Waluya. 2006. Upaya Meningkatkan Penguasaan Konsep Penelitian Kualitatif dalam Metode Pembelajaran. Laporan Teaching Grant: Pend. Matematika Universitas Negeri Semarang.
- Supriyan, Elan. 2015. *Smartphone dan Masyarakat*. Jakarta: Gramedia Press.
- Tirta, I. M. 2009. Analisis Regresi dengan R. Jember: Jember University Press.

Umah, Fachriya. 2014. Penerapan Analisis *Two Step Cluster* dan *Latent Class Cluster* Pada Pengelompokan Data Berskala Campuran Kategorik dan Kontinu. Skripsi. Malang: Universitas Brawijaya.

Vermunt, J. K. dan Magidson, J. 2002. *Latent Class Analysis*. <http://www.statisticalinnovations.com/articles/Latclass.pdf//>. Diakses pada tanggal 18 November 2014.

\_\_\_\_\_. 2002. *Latent Class CLuster Analysis*. <http://members.home.nl/jeroenvermunt/hagenaars2002b.pdf//>. Diakses pada tanggal 18 November 2014.

\_\_\_\_\_. 2002. *Technical Guide for Latent Gold 4.0: Basic and Advance*. Belmont: Statistical Innovations, Inc.

\_\_\_\_\_. 2004. *Latent Class Models for Clustering, A Comparasion of K-Means Clustering*. <http://members.home.nl/jeroenvermunt/cjmr2004.pdf//>. Diakses pada tanggal 18 November 2014.

Welling, Max. 2001. *A Constrained EM Algorithm for Independent Component Analysis*. New York: Jason Willey & Sons Inc.

Yang, Baijian. 2007. *Professional Microsoft Smartphone Programming*. New York: Jason Willey & Sons Inc.

# LAMPIRAN-LAMPIRAN

## LAMPIRAN 1

### Pembuktian Algoritma EM

Prinsip dari algoritma EM dapat dijelaskan menjadi 2 bagian, yaitu sebagai berikut:

#### ***E (Expectation) Step***

E-step dilakukan untuk mencari  $E[\ln[g(y, x, \theta_t)]|x, \hat{\theta}_{t-1}]$ , dimana:  $\hat{\theta}_{t-1}$  adalah taksiran parameter pada iterasi ke  $t - 1$ ,  $t = 1, 2, \dots, T$ ,  $\theta_t$  adalah nilai parameter yang diberikan pada iterasi ke- $t$ , dan  $\theta_0$  adalah suatu nilai taksiran awal yang diberikan (inisialisasi). Dalam *LCCA*, tahapan ini untuk mencari ekspektasi dari:

$$\ln \left[ \sum_{j=1}^J \eta_j g(x_j | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2) \right]$$

untuk setiap kelas ke- $j$  dari peubah laten  $Y$ .

$$E \left[ \ln[g(y, x, \theta_t)] | x, \hat{\theta}_{(t-1)} \right] = E \left[ \ln \left[ \prod_{n=1}^N \left( \eta_j^{(t)} g(x_h | \theta^{(t)}) \right) \right] | x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right]$$

dimana:

$$g(x_h|\theta^{(t)}) = \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}^{(t)}} (1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right] \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}^{(t)}} \right] \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}^{(t)}} \right] \left[ \frac{1}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right] \right]$$

Dengan:

- $\eta_j^{(t)}$  adalah nilai peluang awal kelas laten ke- $j$  pada iterasi ke- $t$ .
- $\pi_{ij}^{(t)}$  adalah peluang objek memberikan respon ya untuk peubah biner pada iterasi ke- $t$ .
- $\pi_{ij(s)}^{(t)}$  adalah peluang objek memberikan respon ke salah satu kategori untuk peubah nominal pada iterasi ke- $t$ .
- $\gamma_{ij(s)} - \gamma_{ij(s-1)}^{(t)}$  adalah peluang objek memberikan respon ke salah satu kategori untuk peubah ordinal pada iterasi ke- $t$ .
- $\mu_{ij}^{(t)}$  adalah nilai rata-rata peubah kontinu pada iterasi ke- $t$ .
- $\sigma_i^{2(t)}$  adalah *varians* pada iterasi ke- $t$ .

$$E \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h|\theta^{(t)}) \right) \right] \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right]$$

$$= \sum_{j=1}^J \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h|\theta^{(t)}) \right) \right] \right] \cdot h \left( j \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right)$$

dimana:

$$\begin{aligned}
& \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2) \right) \right] \\
= & \prod_{h=1}^N \left[ \ln \left[ \eta_j^{(t)} g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2) \right] \right] \\
= & \sum_{h=1}^N \left[ \ln \eta_j^{(t)} + \ln g(x_h | \pi_j) + \ln g(x_h | \pi_{j(s)}) + \ln g(x_h | \gamma_{j(s)} - \gamma_{j(s-1)}) + \ln g(x_h | \mu_j, \sigma_j^2) \right] \\
= & \sum_{h=1}^N \ln \eta_j^{(t)} + \sum_{h=1}^N \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}(t)} (1 - \pi_{ij})^{1-x_{ih}(t)} \right] + \sum_{h=1}^N \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}(t)} \right] \\
+ & \sum_{h=1}^N \ln \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}(t)} \right] + \sum_{h=1}^N \ln \left[ \frac{\exp \left[ \prod_{i=1}^Q - \frac{1}{2\sigma_i^2(t)} (x_{ih} - \mu_{ij}^{(t)})^2 \right]}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right]
\end{aligned}$$

sebut

$$\begin{aligned}
T_{j,h}^{(t-1)} &= h(j | x_h, \hat{\theta}^{(t-1)}) \\
&= h(j | x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)})
\end{aligned}$$

dengan  $j = 1, 2, \dots, J$ , dan  $h = 1, 2, \dots, N$ . Maka, berdasarkan teorema Bayes diperoleh:

$$\begin{aligned}
h(j | x_h) &= \frac{h(Y = j) \cdot g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2)}{\sum_{j=1}^J h(Y = j) \cdot g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2)} \\
&= \frac{\eta_j \cdot g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2)}{\sum_{j=1}^J \eta_j \cdot g(x_h | \pi_j, \pi_{j(s)}, \gamma_{j(s)} - \gamma_{j(s-1)}, \mu_j, \sigma_j^2)} \\
T_{j,h}^{(t-1)} &= \frac{\eta_j \cdot g(x_h | \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)})}{\sum_{j=1}^J \eta_j \cdot g(x_h | \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)})}
\end{aligned}$$

Substitusikan persamaan  $T_{j,h}^{(t-1)}$  ke persamaan awal. Dengan demikian, diperoleh:

$$\begin{aligned}
& E \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h | \theta^{(t)}) \right) \right] \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right] \\
&= \sum_{j=1}^J T_{j,h}^{(t-1)} \cdot \left[ \sum_{h=1}^N \ln \eta_j^{(t)} + \sum_{h=1}^N \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}^{(t)}} (1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right] + \sum_{h=1}^N \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}^{(t)}} \right] \right. \\
&\quad \left. + \sum_{h=1}^N \ln \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}^{(t)}} \right] + \sum_{h=1}^N \ln \left[ \frac{\exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right]}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \right]
\end{aligned}$$

dengan  $t = 1, 2, \dots$ .

### ***M (Maximization) Step***

Setelah melakukan *E-Step*, langkah selanjutnya adalah melakukan *M-step*, dimana akan dicari nilai taksiran untuk  $\eta_j^{(t)}$ ,  $\pi_{ij}^{(t)}$ ,  $\pi_{ij(s)}^{(t)}$ ,  $\gamma_{ij(s)} - \gamma_{ij(s-1)}^{(t)}$ ,  $\mu_{ij}^{(t)}$ , dan  $\sigma_i^{2(t)}$  yang memaksimumkan :

$$\begin{aligned}
& P = E \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h | \theta^{(t)}) \right) \right] \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right] \\
&= \sum_{j=1}^J T_{j,h}^{(t-1)} \cdot \left[ \sum_{h=1}^N \ln \eta_j^{(t)} + \sum_{h=1}^N \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}^{(t)}} (1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right] + \sum_{h=1}^N \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}^{(t)}} \right] \right. \\
&\quad \left. + \sum_{h=1}^N \ln \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}^{(t)}} \right] + \sum_{h=1}^N \ln \left[ \frac{\exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right]}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \right]
\end{aligned}$$

yang didapat dari *E-Step*.

Untuk mencari nilai taksiran dari  $\eta_j^{(t)}$ ,  $\pi_{ij}^{(t)}$ ,  $\pi_{ij(s)}^{(t)}$ ,  $\gamma_{ij(s)} - \gamma_{ij(s-1)}^{(t)}$ ,  $\mu_{ij}^{(t)}$ , dan  $\sigma_i^{2(t)}$  yang memaksimumkan hasil *E-step* diperoleh dengan cara menurunkan-nya secara parsial terhadap parameter  $\eta_j^{(t)}$ ,  $\pi_{ij}^{(t)}$ ,  $\pi_{ij(s)}^{(t)}$ ,  $\gamma_{ij(s)} - \gamma_{ij(s-1)}^{(t)}$ ,  $\mu_{ij}^{(t)}$ , dan  $\sigma_i^{2(t)}$  kemudian disamakan dengan nol. Berikut adalah pembuktian nilai taksiran untuk masing-masing parameter:

### Taksiran Untuk $\eta_j^{(t)}$

Taksiran dari  $\eta_j^{(t)}$  didapat dengan menurunkan persamaan yang didapat pada *E-Step* terhadap  $\eta_j^{(t)}$ . Karena terdapat syarat bahwa  $\sum_{j=1}^J \eta_j = 1$ , maka:

$$\begin{aligned}
I &= E \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h | \theta^{(t)}) \right) \right] \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right] \\
&\quad - \gamma \left[ \sum_{j=1}^J \eta_{(j-1)} \right] \\
&= \sum_{j=1}^J T_{j,h}^{(t-1)} \cdot \left[ \sum_{h=1}^N \ln \eta_j^{(t)} + \sum_{h=1}^N \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}(t)} (1 - \pi_{ij})^{1-x_{ih}(t)} \right] + \sum_{h=1}^N \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}(t)} \right] \right] \\
&\quad + \sum_{h=1}^N \ln \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}(t)} \right] + \sum_{h=1}^N \ln \left[ \frac{\exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right]}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \\
&\quad - \gamma \left[ \sum_{j=1}^J \eta_{(j-1)} \right].
\end{aligned}$$

Taksiran dari  $\eta_j^{(t)}$  didapat dengan menurunkan persamaan  $I$  terhadap  $\eta_j^{(t)}$ , yaitu:

$$\begin{aligned} \frac{\partial I}{\partial \eta_j^{(t)}} &= \frac{\partial}{\partial \eta_j^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \ln \eta_j^{(t)} - \gamma \left[ \sum_{j=1}^J \eta_j^{(t-1)} \right] \right) \\ \frac{\partial I}{\partial \eta_1^{(t)}} &= \frac{\partial}{\partial \eta_1^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{1,h}^{(t-1)} \right) \ln \eta_1^{(t)} - \gamma \left[ \sum_{j=1}^J \eta_j^{(t-1)} \right] \right) \\ &= \frac{\sum_{h=1}^N \left( T_{1,h}^{(t-1)} \right)}{\eta_1^{(t)}} - \gamma = 0 \end{aligned}$$

Dengan demikian didapat:

$$\begin{aligned} \eta_1^{(t)} &= \frac{\sum_{h=1}^N \left( T_{1,h}^{(t-1)} \right)}{\gamma} \\ \eta_2^{(t)} &= \frac{\sum_{h=1}^N \left( T_{2,h}^{(t-1)} \right)}{\gamma} \\ \dots &= \dots \\ \eta_J^{(t)} &= \frac{\sum_{h=1}^N \left( T_{J,h}^{(t-1)} \right)}{\gamma} \end{aligned}$$

Perhatikan bahwa,

$$\sum_{j=1}^J \eta_j = 1 \rightarrow \eta_1 + \eta_2 + \dots + \eta_J = 1$$

Dengan demikian,,

$$\begin{aligned} \frac{\sum_{h=1}^N (T_{1,h}^{(t-1)})}{\gamma} + \dots + \frac{\sum_{h=1}^N (T_{J,h}^{(t-1)})}{\gamma} &= \frac{\sum_{h=1}^N (T_{1,h}^{(t-1)} + T_{2,h}^{(t-1)} + \dots + T_{J,h}^{(t-1)})}{\gamma} \\ &= \frac{\sum_{h=1}^N 1}{\gamma} \\ &= 1 \end{aligned}$$

maka,

$$\frac{N}{\gamma} = 1 \rightarrow \gamma = N$$

Dengan demikian, didapat taksiran untuk  $\eta_j^{(t)}$  adalah:

$$\hat{\eta}_j^{(t)} = \frac{\sum_{h=1}^N (T_{1,h}^{(t-1)})}{N} = \frac{\sum_{h=1}^N \hat{h}(j|x_h)}{N}$$

### Taksiran Untuk $\pi_{ij}^{(t)}$

Taksiran dari  $\pi_{ij}^{(t)}$  didapat dengan menurunkan persamaan yang didapat pada *E-Step* terhadap  $\pi_{ij}^{(t)}$ . Dengan demikian, diperoleh:

$$\begin{aligned}
\frac{\partial P}{\partial \pi_{ij}^{(t)}} &= \frac{\partial}{\partial \pi_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}^{(t)}} (1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right] \right) \\
&= \frac{\partial}{\partial \pi_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \sum_{i=1}^B \ln \pi_{ij}^{x_{ih}^{(t)}} (1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right] \right) \\
&= \frac{\partial}{\partial \pi_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \sum_{i=1}^B \left\{ \ln \pi_{ij}^{x_{ih}^{(t)}} + \ln(1 - \pi_{ij})^{1-x_{ih}^{(t)}} \right\} \right] \right) \\
&= \frac{\partial}{\partial \pi_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \sum_{i=1}^B \left\{ x_{ih} \ln \pi_{ij}^{(t)} + (1 - x_{ih}) \ln(1 - \pi_{ij})^{(t)} \right\} \right] \right) \\
&= \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) \left\{ \frac{x_{ih}}{\pi_{ij}^{(t)}} - \frac{(1 - x_{ih})}{(1 - \pi_{ij})^{(t)}} \right\} \right] = 0
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\pi_{ij}^{(t)}} \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] &= \frac{1}{1 - \pi_{ij}^{(t)}} \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) (1 - x_{ih}) \right] \\
\sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] - \pi_{ij}^{(t)} \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] &= \pi_{ij}^{(t)} \left( \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) - \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] \right) \\
\sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] - \pi_{ij}^{(t)} \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] &= \pi_{ij}^{(t)} \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) - \pi_{ij}^{(t)} \sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] \\
\sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right] &= \pi_{ij}^{(t)} \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \\
\pi_{ij}^{(t)} &= \frac{\sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) x_{ih} \right]}{\sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right)}
\end{aligned}$$

Dengan demikian, didapat taksiran untuk  $\pi_{ij}^{(t)}$  adalah:

$$\hat{\pi}_{ij}^{(t)} = \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih}^{(t)}}{N \hat{\eta}_j^{(t)}}$$

### Taksiran Untuk $\pi_{ij(s)}^{(t)}$

Taksiran dari  $\pi_{ij(s)}^{(t)}$  didapat dengan menurunkan persamaan yang didapat pada *E-Step* terhadap  $\pi_{ij(s)}^{(t)}$ . Karena terdapat syarat bahwa  $\sum_{s=1}^S \pi_{ij(s)} = 1$ , maka:

$$\begin{aligned}
Z &= E \left[ \ln \left[ \prod_{h=1}^N \left( \eta_j^{(t)} g(x_h | \theta^{(t)}) \right) \right] \middle| x_h, \hat{\eta}_j^{(t-1)}, \hat{\pi}_{ij}^{(t-1)}, \hat{\pi}_{j(s)}^{(t-1)}, \hat{\gamma}_{(ij)} - \hat{\gamma}_{ij}^{(t-1)}, \hat{\mu}_{ij}^{(t-1)}, \hat{\sigma}_i^{2(t-1)} \right] \\
&\quad - \gamma \left[ \sum_{s=1}^S \pi_{ij(s-1)} \right] \\
&= \sum_{j=1}^J T_{j,h}^{(t-1)} \cdot \left[ \sum_{h=1}^N \ln \eta_j^{(t)} + \sum_{h=1}^N \ln \left[ \prod_{i=1}^B \pi_{ij}^{x_{ih}(t)} (1 - \pi_{ij})^{1-x_{ih}(t)} \right] + \sum_{h=1}^N \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}(t)} \right] \right] \\
&\quad + \sum_{h=1}^N \ln \left[ \prod_{i=1}^O \prod_{s=1}^R (\gamma_{ij(s)} - \gamma_{ij(s-1)})^{x_{ih(s)}(t)} \right] + \sum_{h=1}^N \ln \left[ \frac{\exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^2(t)} (x_{ih} - \mu_{ij}^{(t)})^2 \right]}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \\
&\quad - \gamma \left[ \sum_{s=1}^S \pi_{ij(s-1)} \right]
\end{aligned}$$

Taksiran dari  $\pi_{ij(s)}^{(t)}$  didapat dengan menurunkan persamaan  $Z$  terhadap  $\pi_{ij(s)}^{(t)}$ , yaitu:

$$\begin{aligned}
\frac{\partial Z}{\partial \pi_{ij(s)}^{(t)}} &= \frac{\partial}{\partial \pi_{ij(s)}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \ln \left[ \prod_{i=1}^L \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}(t)} \right] - \gamma \left[ \sum_{s=1}^S (\pi_{ij(s-1)})^{(t)} \right] \right) \\
&= \frac{\partial}{\partial \pi_{ij(s)}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \sum_{i=1}^L \ln \prod_{s=1}^S (\pi_{ij(s)})^{x_{ih(s)}(t)} \right] - \gamma \left[ \sum_{s=1}^S (\pi_{ij(s-1)})^{(t)} \right] \right) \\
\frac{\partial Z}{\partial \pi_{ij(1)}^{(t)}} &= \frac{\partial}{\partial \pi_{ij(1)}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \sum_{i=1}^L \ln (\pi_{ij(1)})^{x_{ih(1)}(t)} \right] - \gamma \left[ \sum_{s=1}^S (\pi_{ij(s-1)})^{(t)} \right] \right) \\
&= \frac{\sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) x_{ih(1)}}{\pi_{ij(1)}^{(t)}} - \gamma = 0
\end{aligned}$$

Dengan demikian, didapat:

$$\begin{aligned}
 \pi_{ij(1)}^{(t)} &= \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)}) x_{ih(1)}}{\gamma} \\
 \pi_{ij(2)}^{(t)} &= \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)}) x_{ih(2)}}{\gamma} \\
 \dots &= \dots \\
 \pi_{ij(S)}^{(t)} &= \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)}) x_{ih(S)}}{\gamma}
 \end{aligned}$$

Perhatikan bahwa,

$$\sum_{s=1}^S \pi_{ij(s)} = 1 \rightarrow \pi_{ij(1)} + \pi_{ij(2)} + \dots + \pi_{ij(S)} = 1$$

Dengan demikian,,

$$\begin{aligned}
 \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)}) x_{ih(1)}}{\gamma} + \dots + \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)}) x_{ih(S)}}{\gamma} &= \frac{\sum_{h=1}^N (T_{j,h}) (\pi_{ij(1)} + \pi_{ij(2)} + \dots + \pi_{ij(S)})}{\gamma} \\
 &= \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)})}{\gamma} \\
 &= 1
 \end{aligned}$$

Maka,

$$\begin{aligned}
 \frac{\sum_{h=1}^N (T_{j,h}^{(t-1)})}{\gamma} &= 1 \\
 \gamma &= \sum_{h=1}^N (T_{j,h}) \\
 \gamma &= N \hat{\eta}_j
 \end{aligned}$$

Dengan demikian, didapat taksiran untuk  $\pi_{ij(s)}^{(t)}$  adalah:

$$\hat{\pi}_{ij(s)}^{(t)} = \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih(s)}^{(t)}}{N \hat{\eta}_j^{(t)}}$$

**Taksiran Untuk  $(\gamma_{ij(s)} - \gamma_{ij(s)})^{(t)}$**

Karena peubah nominal dan peubah ordinal menggunakan model distribusi yang sama, maka didapat taksiran untuk parameter pada peubah ordinal, yaitu:

$$\hat{\pi}_{ij(s)}^{(t)} = \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih(s)}^{(t)}}{N \hat{\eta}_j^{(t)}}$$

Untuk menjaga tingkat urutan dari  $\pi_{ij(s)}^{(t)}$ , maka  $\pi_{ij(s)}^{(t)}$  diubah menjadi  $\gamma_{ij(s)} - \gamma_{ij(s-1)}^{(t)}$ , dengan

$$\gamma_{ij(s)}^{(t)} = \pi_{ij(1)}^{(t)} + \pi_{ij(2)}^{(t)} + \dots + \pi_{ij(s)}^{(t)}$$

Dengan demikian, didapat:

$$\begin{aligned} \hat{\gamma}_{ij(s)} - \hat{\gamma}_{ij(s-1)}^{(t)} &= \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih(s)}^{(t)}}{N \hat{\eta}_j^{(t)}} \\ \hat{\gamma}_{ij(s)}^{(t)} &= \hat{\gamma}_{ij(s-1)}^{(t)} + \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih(s)}^{(t)}}{N \hat{\eta}_j^{(t)}} \end{aligned}$$

**Taksiran Untuk  $\mu_{ij}^{(t)}$**

Taksiran dari  $\mu_{ij}^{(t)}$  didapat dengan menurunkan persamaan yang didapat pada *E-Step* terhadap  $\mu_{ij}^{(t)}$ . Dengan demikian, diperoleh:

$$\begin{aligned}
\frac{\partial P}{\partial \mu_{ij}^{(t)}} &= \frac{\partial}{\partial \mu_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N (T_{j,h}^{(t-1)}) \ln \left[ \frac{1}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right] \right) \\
&= \frac{\partial}{\partial \mu_{ij}^{(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N (T_{j,h}^{(t-1)}) \left[ - \left( \sum_{i=1}^Q \frac{(x_{ih} - \mu_{ij}^{(t)})^2}{2\sigma_i^{2(t)}} \right) - \left( \frac{Q}{2} \ln(2\pi) \right) - \left( \sum_{i=1}^Q \ln \sigma_i^{2(t)} \right) \right] \right) \\
&= \sum_{h=1}^N (T_{j,h}^{(t-1)}) \left[ - \left( \frac{2(-1)(x_{ih} - \mu_{ij}^{(t)})}{2\sigma_i^{2(t)}} \right) \right] \\
&= \sum_{h=1}^N (T_{j,h}^{(t-1)}) \left[ \left( \frac{(x_{ih} - \mu_{ij}^{(t)})}{\sigma_i^{2(t)}} \right) \right] = 0
\end{aligned}$$

Asumsikan bahwa  $\sigma_i^{2(t)}$  berhingga, Dengan demikian, didapat taksiran untuk  $\mu_{ij}^{(t)}$ , yaitu :

$$\begin{aligned}
\sum_{h=1}^N (T_{j,h}^{(t-1)}) (x_{ih} - \mu_{ij}^{(t)}) &= 0 \\
\sum_{h=1}^N \left[ (x_{ih} T_{j,h}^{(t-1)}) - (\mu_{ij}^{(t)} T_{j,h}^{(t-1)}) \right] &= 0 \\
\sum_{h=1}^N [x_{ih} T_{j,h}^{(t-1)}] &= \mu_{ij}^{(t)} \sum_{h=1}^N T_{j,h}^{(t-1)} \\
\hat{\mu}_{ij}^{(t)} &= \frac{\sum_{h=1}^N [x_{ih} T_{j,h}^{(t-1)}]}{\sum_{h=1}^N T_{j,h}^{(t-1)}} \\
\hat{\mu}_{ij}^{(t)} &= \frac{\sum_{h=1}^N \hat{h}(j|x_h) x_{ih}^{(t)}}{N \hat{\eta}_j^{(t)}}
\end{aligned}$$

### Taksiran Untuk $\sigma_i^{2(t)}$

Taksiran dari  $\sigma_i^{2(t)}$  didapat dengan menurunkan persamaan yang didapat pada *E-Step* terhadap  $\sigma_i^{2(t)}$ . Dengan demikian, diperoleh:

$$\begin{aligned}
\frac{\partial P}{\partial \sigma_i^{2(t)}} &= \frac{\partial}{\partial \sigma_i^{2(t)}} \left( \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \ln \left[ \frac{1}{\sqrt{2\pi} \prod_{i=1}^Q \sigma_i^{(t)}} \right] \exp \left[ \prod_{i=1}^Q -\frac{1}{2\sigma_i^{2(t)}} (x_{ih} - \mu_{ij}^{(t)})^2 \right] \right) \\
&= \frac{\partial}{\partial \sigma_i^{2(t)}} \sum_{j=1}^J \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ - \left( \sum_{i=1}^Q \frac{(x_{ih} - \mu_{ij}^{(t)})^2}{2\sigma_i^{2(t)}} \right) - \left( \frac{Q}{2} \ln(2\pi) \right) - \left( \sum_{i=1}^Q \ln \sigma_i^{2(t)} \right) \right] \\
&= \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \left( \frac{(x_{ih} - \mu_{ij}^{(t)})^2}{2(\sigma_i^{2(t)})^2} - \frac{1}{2\sigma_i^{2(t)}} \right) \right] \\
&= \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ \left( \frac{(x_{ih} - \mu_{ij}^{(t)})^2 - \sigma_i^{2(t)}}{2(\sigma_i^{2(t)})^2} \right) \right] = 0
\end{aligned}$$

Asumsikan  $\sigma_i^{2(t)}$  berhingga. Dengan demikian,:

$$\begin{aligned}
\sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ (x_{ih} - \mu_{ij}^{(t)})^2 - \sigma_i^{2(t)} \right] &= 0 \\
\sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right) \left[ (x_{ih} - \mu_{ij}^{(t)})^2 \right] &= \sigma_i^{2(t)} \sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right)
\end{aligned}$$

Jadi, didapatkan taksiran untuk  $\sigma_i^{2(t)}$  yaitu:

$$\hat{\sigma}_i^{2(t)} = \frac{\sum_{h=1}^N \left[ \left( T_{j,h}^{(t-1)} \right) \left[ (x_{ih} - \mu_{ij}^{(t)})^2 \right] \right]}{\sum_{h=1}^N \left( T_{j,h}^{(t-1)} \right)} = \frac{\sum_{h=1}^N \sum_{j=1}^J (x_{ih} - \hat{\mu}_{ij}^{(t)})^2 \hat{h}(j|x_h)}{\sum_{h=1}^N \sum_{j=1}^J \hat{h}(j|x_h)}$$

Proses *E-Step* dan *M-Step* ini dilakukan secara iteratif sampai didapatkan suatu nilai taksiran yang konvergen atau didapatkan  $|\hat{\theta}_t - \hat{\theta}_{(t-1)}|$  mendekati nol. Di dalam *software* LatentGOLD 4.0 iterasi akan berhenti jika pergantian dalam *log-posterior* kurang dari  $10^{-12}$ .

## LAMPIRAN 2

Statistik Deskriptif Peubah yang Digunakan

Statistik Deskriptif Peubah Kategorik yang Digunakan

**Jaringan yang digunakan (X1)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	EDGE	4	2,1	2,1	2,1
	HSDPA	96	49,5	49,5	51,5
	LTE	94	48,5	48,5	100,0
	Total	194	100,0	100,0	

**Kemampuan Dual SIM (X2)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tidak	116	59,8	59,8	59,8
	Ya	78	40,2	40,2	100,0
	Total	194	100,0	100,0	

**Jenis SIM yang digunakan (X3)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mini SIM	55	28,4	28,4	28,4
	Micro SIM	118	60,8	60,8	89,2
	Nano SIM	21	10,8	10,8	100,0
	Total	194	100,0	100,0	

**Lebar layar (X4)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<4,0	20	10,3	10,3	10,3
	4,0-4,99	72	37,1	37,1	47,4
	5-5,99	65	33,5	33,5	80,9
	6-6,99	11	5,7	5,7	86,6
	>=7	26	13,4	13,4	100,0
	Total	194	100,0	100,0	

**Jenis sistem operasi (X5)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Android	156	80,4	80,4	80,4
	Windows Phone	22	11,3	11,3	91,8
	IOS	11	5,7	5,7	97,4
	Blackberry	5	2,6	2,6	100,0
	Total	194	100,0	100,0	

**Jenis processor (X6)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Single Core	19	9,8	9,8	9,8
	Dual Core	73	37,6	37,6	47,4
	Quad Core	96	49,5	49,5	96,9
	Octa Core	6	3,1	3,1	100,0
	Total	194	100,0	100,0	

**Ketersediaan memori eksternal (X7)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tidak	33	17,0	17,0	17,0
	Ada	161	83,0	83,0	100,0
	Total	194	100,0	100,0	

**Memori internal yang tersedia (X8)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4GB	46	23,7	23,7	23,7
	8GB	54	27,8	27,8	51,5
	16GB	41	21,1	21,1	72,7
	32GB	40	20,6	20,6	93,3
	64GB	9	4,6	4,6	97,9
	128GB	4	2,1	2,1	100,0
	Total	194	100,0	100,0	

**Kapasitas RAM yang tersedia (X9)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	512MB	37	19,1	19,1	19,1
	768MB	7	3,6	3,6	22,7
	1GB	72	37,1	37,1	59,8
	1,5GB	14	7,2	7,2	67,0
	2GB	48	24,7	24,7	91,8
	3GB	14	7,2	7,2	99,0
	4GB	2	1,0	1,0	100,0
	Total	194	100,0	100,0	

**Kamera utama (X10)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<5,0 MP	29	14,9	14,9	14,9
	5,0-7,99	58	29,9	29,9	44,8
	8,0-12,99	61	31,4	31,4	76,3
	>=13	46	23,7	23,7	100,0
	Total	194	100,0	100,0	

**Ketersediaan flash (X11)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tidak	47	24,2	24,2	24,2
	Ya	147	75,8	75,8	100,0
	Total	194	100,0	100,0	

**Kamera depan (X12)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tidak Ada	26	13,4	13,4	13,4
	VGA-1,99 MP	96	49,5	49,5	62,9
	2,00-4,99 MP	49	25,3	25,3	88,1
	>=5,00 MP	23	11,9	11,9	100,0
	Total	194	100,0	100,0	

**Kapasitas baterai (X15)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<=2000mAh	76	39,2	39,2	39,2
	2001-3000 mAh	72	37,1	37,1	76,3
	3001-4000 mAh	30	15,5	15,5	91,8
	>4000 mAh	16	8,2	8,2	100,0
	Total	194	100,0	100,0	

## Statistik Deskriptif Peubah Kontinu yang Digunakan

Statistics				Statistics			
		Harga produk pada Mei 2015 (X13)	Popularitas pada Mei 2015 (X14)			Ln Harga	Ln Popularity
N	Valid	194	194	N	Valid	194	194
	Missing	0	0		Missing	0	0
Mean		3,3825206	8,4675258	Mean		,9918	1,5226010
Median		2,6500000	4,3000000	Median		,9746	1,4586150
Mode		3,60000	1,70000	Mode		1,28	,53063
Std. Deviation		2,45938049	10,53130183	Std. Deviation		,67489	1,12965116
Variance		6,049	110,908	Variance		,455	1,276
Range		13,17600	64,70000	Range		3,10	5,37836
Minimum		,62400	,30000	Minimum		-,47	-1,20397
Maximum		13,80000	65,00000	Maximum		2,62	4,17439
Sum		656,20900	1642,70000	Sum		192,40	295,38459

## Nilai Korelasi Peubah Kontinu yang Digunakan

Correlations			
		Ln Harga	Ln Popularity
Ln Harga	Pearson Correlation	1	,268**
	Sig. (2-tailed)		,000
	N	194	194
Ln Popularity	Pearson Correlation	,268**	1
	Sig. (2-tailed)	,000	
	N	194	194

\*\* . Correlation is significant at the 0.01 level (2-tailed).

### LAMPIRAN 3

Nilai BVR Model 5 Gerombol Tanpa *Direct Effect*

Indicators	Jaringan	DualSIM	JenisSIM	UkuranLayar	Platform	CPU	Eksternal
Jaringan	.						
DualSIM	8,5996	.					
JenisSIM	4,0824	0,0570	.				
UkuranLayar	0,0076	2,5787	0,0587	.			
Platform	1,3965	2,5044	5,1009	5,4981	.		
CPU	1,0509	7,3123	0,3587	3,8998	7,0675	.	
Eksternal	0,3901	0,0121	1,4755	0,0450	1,9535	0,1012	.
Internal	2,9145	1,1566	5,5102	0,0793	3,7903	0,0294	8,9841
RAM	1,9513	0,1765	0,4789	1,3435	0,4787	3,0335	0,0005
KameraUtama	1,4089	0,1407	0,1706	0,2656	1,9925	2,5744	1,2297
Flash	0,0071	0,6707	0,6379	0,0265	0,2506	0,6026	0,0098
KameraDepan	0,2080	6,5546	0,1452	0,5634	0,0958	6,5966	0,0200
Kapasitas	1,1791	0,3978	0,0800	7,2880	0,4362	0,5810	0,0045

Indicators	Internal	RAM	KameraUtama	Flash	KameraDepan	Kapasitas
Jaringan						
DualSIM						
JenisSIM						
UkuranLayar						
Platform						
CPU						
Eksternal						
Internal	.					
RAM	4,3762	.				
KameraUtama	2,7493	2,4271	.			
Flash	0,0572	0,0544	0,6254	.		
KameraDepan	0,3390	1,0669	1,0151	1,0453	.	
Kapasitas	2,4744	2,4028	0,1067	0,0018	0,8655	.

## LAMPIRAN 4

Nilai BVR Model 5 Gerombol dengan 6 *Direct Effect*

Indicators	Jaringan	DualSIM	JenisSIM	UkuranLayar	Platform	CPU	Eksternal
Jaringan	.	.	.	.	.	.	.
DualSIM	5,5426	.	.	.	.	.	.
JenisSIM	1,7920	0,1539	.	.	.	.	.
UkuranLayar	0,0001	1,0316	0,0351	.	.	.	.
Platform	1,0416	2,8350	3,5966	0,9496	.	.	.
CPU	1,2886	0,0000	0,2948	0,0000	0,0000	.	.
Eksternal	1,1924	1,3120	1,2267	0,9238	2,7280	1,2081	.
Internal	0,9233	2,0671	2,5522	0,0333	3,0329	0,0007	0,0000
RAM	0,9842	2,0019	0,4516	4,4098	0,7556	6,4175	0,0585
KameraUtama	0,6740	0,0170	0,1289	0,1596	4,3958	3,7924	1,5970
Flash	0,0179	0,0824	0,5837	1,1656	0,3490	0,8224	0,0007
KameraDepan	0,6291	0,0000	0,3331	0,6234	0,1524	2,8011	0,1501
Kapasitas	0,9569	0,3295	0,0705	0,0000	0,0696	0,0000	1,2996

Indicators	Internal	RAM	KameraUtama	Flash	KameraDepan	Kapasitas
Jaringan	.	.	.	.	.	.
DualSIM	.	.	.	.	.	.
JenisSIM	.	.	.	.	.	.
UkuranLayar	.	.	.	.	.	.
Platform	.	.	.	.	.	.
CPU	.	.	.	.	.	.
Eksternal	.	.	.	.	.	.
Internal	.	.	.	.	.	.
RAM	5,2492	.	.	.	.	.
KameraUtama	3,2924	3,2434	.	.	.	.
Flash	0,0872	0,1604	0,7098	.	.	.
KameraDepan	0,2841	1,6262	0,9174	0,2605	.	.
Kapasitas	1,4731	6,1582	0,0090	0,8798	0,3072	.

## LAMPIRAN 5

Nilai Sebaran Peubah Di Setiap Gerombol Model 5 Gerombol dengan 6 *Direct Effect*

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
<b>Overall</b>	0,3386	0,3264	0,1892	0,0986	0,0472
<b>Indicators</b>					
<b>Jaringan</b>					
<b>EDGE</b>	0,0000	0,0000	0,7491	0,2509	0,0000
<b>HSDPA</b>	0,5042	0,0501	0,3514	0,0943	0,0000
<b>LTE</b>	0,1840	0,6222	0,0000	0,0965	0,0973
<b>DualSIM</b>					
<b>Tidak</b>	0,2433	0,4355	0,0863	0,1561	0,0788
<b>Ya</b>	0,4805	0,1639	0,3427	0,0129	0,0000
<b>JenisSIM</b>					
<b>Mini SIM</b>	0,4823	0,0182	0,4081	0,0914	0,0000
<b>Micro SIM</b>	0,3323	0,4270	0,1211	0,1196	0,0000
<b>Nano SIM</b>	0,0000	0,5668	0,0000	0,0000	0,4332
<b>UkuranLayar</b>					
<b>&lt;4,0</b>	0,1520	0,0996	0,7484	0,0000	0,0000
<b>4,0-4,99</b>	0,4776	0,1638	0,3021	0,0000	0,0565
<b>5-5,99</b>	0,3800	0,6043	0,0000	0,0000	0,0157
<b>6-6,99</b>	0,3293	0,6684	0,0000	0,0023	0,0000
<b>&gt;=7</b>	0,0000	0,1144	0,0000	0,7301	0,1555
<b>Platform</b>					
<b>Android</b>	0,3279	0,3593	0,1967	0,1161	0,0000
<b>Windows Phone</b>	0,5266	0,1509	0,2756	0,0468	0,0000
<b>IOS</b>	0,1790	0,0000	0,0000	0,0000	0,8210
<b>Blackberry</b>	0,2010	0,7990	0,0000	0,0000	0,0000
<b>CPU</b>					
<b>Single Core</b>	0,1797	0,0000	0,8203	0,0000	0,0000
<b>Dual Core</b>	0,4614	0,1247	0,2197	0,0829	0,1113
<b>Quad Core</b>	0,2872	0,5235	0,0424	0,1362	0,0106
<b>Octa Core</b>	0,1670	0,6660	0,1670	0,0000	0,0000
<b>Eksternal</b>					
<b>Tidak</b>	0,2316	0,4616	0,0000	0,0304	0,2764
<b>Ada</b>	0,3606	0,2986	0,2282	0,1126	0,0000

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
<b>Overall</b>	0,3386	0,3264	0,1892	0,0986	0,0472
<b>Internal</b>					
<b>4GB</b>	0,3760	0,0000	0,6240	0,0000	0,0000
<b>8GB</b>	0,6399	0,1363	0,1488	0,0750	0,0000
<b>16GB</b>	0,2039	0,5753	0,0000	0,2208	0,0000
<b>32GB</b>	0,1134	0,7103	0,0000	0,1509	0,0255
<b>64GB</b>	0,1102	0,4407	0,0000	0,0000	0,4491
<b>128GB</b>	0,0000	0,0000	0,0000	0,0000	1,0000
<b>RAM</b>					
<b>512MB</b>	0,2426	0,0000	0,7574	0,0000	0,0000
<b>768MB</b>	0,5108	0,0000	0,4892	0,0000	0,0000
<b>1GB</b>	0,6504	0,0376	0,0732	0,1259	0,1129
<b>1,5GB</b>	0,2802	0,3608	0,0000	0,3589	0,0000
<b>2GB</b>	0,0486	0,8254	0,0000	0,1048	0,0212
<b>3GB</b>	0,0000	1,0000	0,0000	0,0000	0,0000
<b>4GB</b>	0,0000	1,0000	0,0000	0,0000	0,0000
<b>KameraUtama</b>					
<b>&lt;5,0 MP</b>	0,0000	0,0000	0,7231	0,2769	0,0000
<b>5,0-7,99</b>	0,4689	0,0171	0,2705	0,1909	0,0526
<b>8,0-12,99</b>	0,5881	0,3118	0,0000	0,0000	0,1001
<b>&gt;=13</b>	0,0566	0,9434	0,0000	0,0000	0,0000
<b>Flash</b>					
<b>Tidak</b>	0,0005	0,0000	0,5296	0,3836	0,0862
<b>Ya</b>	0,4472	0,4313	0,0799	0,0070	0,0346
<b>KameraDepan</b>					
<b>Tidak Ada</b>	0,0036	0,0000	0,9577	0,0386	0,0000
<b>VGA-1,99 MP</b>	0,5349	0,1212	0,1123	0,1363	0,0953
<b>2,00-4,99 MP</b>	0,2309	0,6458	0,0206	0,1027	0,0000
<b>&gt;=5,00 MP</b>	0,1259	0,8741	0,0000	0,0000	0,0000
<b>Inharga</b>					
<b>-0,472 - 0,372</b>	0,0979	0,0000	0,8209	0,0812	0,0000
<b>0,405 - 0,742</b>	0,7036	0,0246	0,1496	0,1221	0,0000
<b>0,777 - 1,194</b>	0,6606	0,2551	0,0051	0,0792	0,0000
<b>1,209 - 1,558</b>	0,1922	0,6530	0,0000	0,1548	0,0000
<b>1,609 - 2,625</b>	0,0188	0,6988	0,0000	0,0514	0,2310
<b>InPopularity</b>					
<b>-1,204 - 0,531</b>	0,2997	0,0767	0,3477	0,2760	0,0000
<b>0,588 - 1,131</b>	0,2584	0,2886	0,2944	0,1057	0,0530
<b>1,163 - 1,723</b>	0,3225	0,4523	0,1188	0,0532	0,0531
<b>1,740 - 2,398</b>	0,4924	0,2426	0,1860	0,0526	0,0263
<b>2,485 - 4,174</b>	0,3246	0,5739	0,0009	0,0000	0,1006
<b>Kapasitas</b>					
<b>&lt;=2000mAh</b>	0,3902	0,0732	0,4830	0,0000	0,0535
<b>2001-3000 mAh</b>	0,4269	0,5589	0,0000	0,0000	0,0142
<b>3001-4000 mAh</b>	0,1780	0,4862	0,0000	0,3358	0,0000
<b>&gt;4000 mAh</b>	0,0000	0,1859	0,0000	0,5616	0,2525

## LAMPIRAN 6

Nilai *Direct Effect* Model 5 Gerombol dengan 6 *Direct Effect*

Direct Effects						
CPU	DualSIM					
	Tidak	Ya	Wald	p-value		
	-0,8718	0,8718	20,7540	5,2e-6		
CPU	UkuranLayar		Wald	p-value		
	0,9727	19,6831	9,1e-6			
CPU	Platform					
	Android	Windows Phone	IOS	Blackberry	Wald	p-value
	1,2966	1,2156	-1,0231	-1,4891	13,0177	0,0046
Internal	Eksternal					
	Tidak	Ada	Wald	p-value		
	1,0323	-1,0323	22,9606	1,7e-6		
KameraDepan	DualSIM					
	Tidak	Ya	Wald	p-value		
	-1,1210	1,1210	15,2939	9,2e-5		
Kapasitas	UkuranLayar		Wald	p-value		
	2,4953	48,0175	4,2e-12			

## LAMPIRAN 7

Nilai Iterasi Algoritma EM dan NR Model 5 Gerombol dengan 6 *Direct Effect*

Iteration		log-posterior	log-likelihood	criterion
	0	-4189,6745741082	-4168,2069518368	
EM	5	-2392,6283764662	-2370,5555734315	56,69915072
EM	10	-2284,1481390071	-2257,6092986435	7,43645330
EM	15	-2243,3560196057	-2214,7929973568	12,01536167
EM	20	-2217,1561127826	-2186,7193854913	11,03875957
EM	25	-2211,2424597804	-2179,8923695069	3,32402970
EM	30	-2208,6486435028	-2176,7270855080	2,65684854
EM	35	-2207,1656222817	-2174,8095337201	1,81019744
EM	40	-2206,2736425687	-2173,5759119591	9,80160639
EM	45	-2205,7288035867	-2172,7580681631	0,92427174
EM	50	-2205,3890662338	-2172,1998059291	0,64227771
EM	55	-2205,1715932236	-2171,8069737894	0,51044640
EM	60	-2205,0291276025	-2171,5229216106	0,45301108
EM	65	-2204,9340328015	-2171,3126001496	0,50852451
EM	70	-2204,8695923326	-2171,1536230818	1,71278729
EM	75	-2204,8253836337	-2171,0312710491	0,32726948
EM	80	-2204,7947417983	-2170,9356105975	0,22048574
EM	85	-2204,7733161834	-2170,8597801668	0,17381252
EM	90	-2204,7582191565	-2170,7989400893	0,14503825
EM	95	-2204,7475078921	-2170,7496115361	0,12697612
EM	100	-2204,7398603526	-2170,7092499919	0,11468891
EM	105	-2204,7343682009	-2170,6759640646	0,10626689
EM	110	-2204,7304020742	-2170,6483265672	0,10130062
EM	115	-2204,7275226475	-2170,6252454167	0,10001750
EM	120	-2204,7254212029	-2170,6058740207	0,10358908
EM	125	-2204,7238794944	-2170,5895481519	0,11540781
EM	130	-2204,7227423917	-2170,5757408574	0,14669391
EM	135	-2204,7218990784	-2170,5640298145	0,26004208
EM	140	-2204,7212700313	-2170,5540733949	1,81141371
EM	145	-2204,7207979425	-2170,5455928972	0,18786714
EM	150	-2204,7204413498	-2170,5383592023	0,10116973
EM	155	-2204,7201701421	-2170,5321826405	0,06911777
EM	160	-2204,7199623725	-2170,5269052139	0,05200825
EM	165	-2204,7198019860	-2170,5223945674	0,04117716
EM	170	-2204,7196771952	-2170,5185392703	0,03364376
EM	175	-2204,7195793119	-2170,5152450903	0,02813004
EM	180	-2204,7195019064	-2170,5124320259	0,02388349
EM	185	-2204,7194401991	-2170,5100319235	0,02049870
EM	190	-2204,7193906193	-2170,5079865491	0,01773778
EM	195	-2204,7193504850	-2170,5062460171	0,01543629
EM	200	-2204,7193177695	-2170,5047675006	0,01349151
EM	205	-2204,7192909308	-2170,5035141664	0,01183133
EM	210	-2204,7192687874	-2170,5024542892	0,01040189
EM	212	-2204,7192610328	-2170,5020782464	0,00988527
Newton	1	-2204,7191552486	-2170,4980433943	0,10211847
Newton	2	-2204,7191552419	-2170,4979560353	0,00079781
Newton	3	-2204,7191552419	-2170,4979560336	0,00000002
Newton	4	-2204,7191552419	-2170,4979560291	0,00000005
Newton	5	-2204,7191552419	-2170,4979560291	0,00000000
Standard errors				
Preparing output				
Wall clock time=116,5427 CPU=113,4127				

## LAMPIRAN 8

Tabel 1: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
1	Apple	iPhone 4	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
2	Apple	iPhone 4s	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
3	Apple	iPhone 5	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
4	Apple	iPad Air	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
5	Apple	iPhone 5c	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
6	Apple	iPad Mini 2	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
7	Apple	iPhone 5s	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
8	Apple	iPad Mini 3	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
9	Apple	iPad Air 2	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
10	Apple	iPhone 6	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
11	Apple	iPhone 6 Plus	0,0000	0,0000	0,0000	0,0000	<b>1,0000</b>
12	ASUS	Zenfone 4	0,0018	0,0000	<b>0,9982</b>	0,0000	0,0000
13	ASUS	Zenfone C	<b>0,9990</b>	0,0000	0,0010	0,0000	0,0000
14	ASUS	Fonepad 7	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
15	ASUS	Zenfone 4s	<b>0,9998</b>	0,0000	0,0002	0,0000	0,0000
16	ASUS	Google Nexus	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
17	ASUS	Zenfone 5	<b>0,9928</b>	0,0072	0,0000	0,0000	0,0000
18	ASUS	Zenfone 6	0,3286	<b>0,6714</b>	0,0000	0,0000	0,0000
19	ASUS	Zenfone Zoom	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
20	ASUS	Fonepad 8	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
21	ASUS	Zenfone 2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
22	ASUS	Padfone S	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
23	ASUS	Zenfone 2s	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
24	BlackBerry	Z3	<b>0,9995</b>	0,0005	0,0000	0,0000	0,0000
25	BlackBerry	Q5	0,0055	<b>0,9945</b>	0,0000	0,0000	0,0000
26	BlackBerry	Z10	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
27	BlackBerry	Q10	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
28	BlackBerry	Z30	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
29	Lenovo	A316	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
30	Lenovo	A390	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
31	Lenovo	A369	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
32	Lenovo	Muszik A319	0,0001	0,0000	<b>0,9999</b>	0,0000	0,0000
33	Lenovo	A516	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
34	Lenovo	A526	0,0004	0,0000	<b>0,9996</b>	0,0000	0,0000
35	Lenovo	A3300	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000

Tabel 2: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
36	Lenovo	A859	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
37	Lenovo	A3500	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
38	Lenovo	A536	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
39	Lenovo	A6000	<b>0,9996</b>	0,0004	0,0000	0,0000	0,0000
40	Lenovo	A850	<b>0,9998</b>	0,0000	0,0002	0,0000	0,0000
41	Lenovo	S890	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
42	Lenovo	A706	<b>0,9459</b>	0,0000	0,0541	0,0000	0,0000
43	Lenovo	K3 Note	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
44	Lenovo	S920	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
45	Lenovo	P780	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
46	Lenovo	A5500	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
47	Lenovo	S820	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
48	Lenovo	S930	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
49	Lenovo	Golden Warrior	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
50	Lenovo	A7000	0,0027	<b>0,9973</b>	0,0000	0,0000	0,0000
51	Lenovo	S860	<b>0,9978</b>	0,0022	0,0000	0,0000	0,0000
52	Lenovo	Vibe X	0,0061	<b>0,9939</b>	0,0000	0,0000	0,0000
53	Lenovo	K900	0,0021	<b>0,9979</b>	0,0000	0,0000	0,0000
54	Lenovo	P70	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
55	Lenovo	Vibe X2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
56	Lenovo	Vibe Z	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
57	Lenovo	Livo S90	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
58	Lenovo	P90	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
59	Lenovo	S856	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
60	Lenovo	S939	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
61	Lenovo	Vibe Shot	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
62	Lenovo	Vibe Z2 Pro	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
63	LG	Optimus L1 2	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
64	LG	A690	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
65	LG	L40 Dual	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
66	LG	Optimus L3 II	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
67	LG	L60 Dual	0,0079	0,0000	<b>0,9921</b>	0,0000	0,0000
68	LG	Optimus L4	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
69	LG	L80 Dual	<b>0,9986</b>	0,0000	0,0014	0,0000	0,0000
70	LG	Optimus L5	0,0044	0,0000	<b>0,9956</b>	0,0000	0,0000
71	LG	L Fino	<b>0,9999</b>	0,0000	0,0001	0,0000	0,0000
72	LG	Optimus L7	<b>0,8088</b>	0,0000	0,1912	0,0000	0,0000
73	LG	Optimus L7 II	<b>0,9972</b>	0,0000	0,0028	0,0000	0,0000

Tabel 3: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
74	LG	G Pro Lite	<b>0,9999</b>	0,0001	0,0000	0,0000	0,0000
75	LG	L Bello	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
76	LG	G2 Mini	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
77	LG	Optimus L9 II	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
78	LG	G3 Stylus	<b>0,9998</b>	0,0002	0,0000	0,0000	0,0000
79	LG	G Pad 8,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
80	LG	Nexus 5	0,0002	<b>0,9998</b>	0,0000	0,0000	0,0000
81	LG	Optimus G	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
82	LG	G2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
83	LG	Optimus G Pro	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
84	LG	G Pro 2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
85	LG	G3	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
86	LG	G Flex	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
87	LG	G Flex 2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
88	Microsoft	Lumia 435 DS	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
89	Microsoft	Lumia 532 DS	0,0152	0,0000	<b>0,9848</b>	0,0000	0,0000
90	Microsoft	Lumia 535 DS	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
91	Nokia	X DS	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
92	Nokia	X+	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
93	Nokia	Lumia 520	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
94	Nokia	Lumia 530 DS	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
95	Nokia	Lumia 620	<b>0,9213</b>	0,0000	0,0787	0,0000	0,0000
96	Nokia	X2 DS	<b>0,8390</b>	0,0000	0,1610	0,0000	0,0000
97	Nokia	Lumia 625	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
98	Nokia	XL DS	<b>0,9993</b>	0,0000	0,0007	0,0000	0,0000
99	Nokia	Lumia 525	0,0083	0,0000	<b>0,9917</b>	0,0000	0,0000
100	Nokia	Lumia 900	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
101	Nokia	Lumia 630 DS	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
102	Nokia	Lumia 820	<b>0,9999</b>	0,0001	0,0000	0,0000	0,0000
103	Nokia	Lumia 720	<b>0,9993</b>	0,0000	0,0007	0,0000	0,0000
104	Nokia	Lumia 730 DS	0,9999	0,0001	0,0000	0,0000	0,0000
105	Nokia	Lumia 1320	<b>0,9754</b>	0,0010	0,0000	0,0236	0,0000
106	Nokia	Lumia 925	<b>0,9432</b>	0,0568	0,0000	0,0000	0,0000
107	Nokia	Lumia 2520	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
108	Nokia	Lumia 920	<b>0,7343</b>	0,2655	0,0000	0,0000	0,0003
109	Nokia	Lumia 1020	0,0001	<b>0,9999</b>	0,0000	0,0000	0,0000
110	Nokia	Lumia 930	0,0000	<b>0,9995</b>	0,0000	0,0000	0,0005
111	Nokia	Lumia 1520	0,0000	<b>0,9999</b>	0,0000	0,0000	0,0001

Tabel 4: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
112	Samsung	Galaxy Young 2	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
113	Samsung	Galaxy star	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
114	Samsung	Galaxy Chat	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
115	Samsung	Galaxy Star Deluxe	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
116	Samsung	Galaxy Pocket Neo	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
117	Samsung	Galaxy young S6310	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
118	Samsung	Galaxy Star Pro Duos	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
119	Samsung	Galaxy V	0,0002	0,0000	<b>0,9998</b>	0,0000	0,0000
120	Samsung	Galaxy Fame	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
121	Samsung	Galaxy Ace 4	<b>0,5640</b>	0,0000	0,4360	0,0000	0,0000
122	Samsung	Galaxy Core	<b>0,9485</b>	0,0000	0,0515	0,0000	0,0000
123	Samsung	Galaxy J1	<b>0,9951</b>	0,0000	0,0049	0,0000	0,0000
124	Samsung	Galaxy Ace 3	<b>0,9912</b>	0,0000	0,0088	0,0000	0,0000
125	Samsung	Galaxy Tab 2 7,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
126	Samsung	Galaxy Core II	<b>0,9688</b>	0,0000	0,0312	0,0000	0,0000
127	Samsung	Galaxy S2	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
128	Samsung	Galaxy Tab 3V	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
129	Samsung	Galaxy Tab 3 7 Lite	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
130	Samsung	Galaxy S duos 2	<b>0,6144</b>	0,0000	0,3856	0,0000	0,0000
131	Samsung	Galaxy Ativ I930	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
132	Samsung	Galaxy Grand Neo Plus	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
133	Samsung	Galaxy Ativ S	<b>0,9995</b>	0,0005	0,0000	0,0000	0,0000
134	Samsung	Grand Neo	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
135	Samsung	Galaxy Grand Prime	<b>0,8830</b>	0,1170	0,0000	0,0000	0,0000
136	Samsung	Galaxy S3 Mini	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
137	Samsung	Grand 2	<b>0,9958</b>	0,0042	0,0000	0,0000	0,0000
138	Samsung	Galaxy Tab 3 7,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
139	Samsung	Galaxy S3	<b>0,8577</b>	0,1423	0,0000	0,0000	0,0000
140	Samsung	Galaxy Grand	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
141	Samsung	Galaxy Mega 5,8	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
142	Samsung	Galaxy A3	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
143	Samsung	Galaxy E5 Duos	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
144	Samsung	Galaxy Mega 2	0,0014	<b>0,9986</b>	0,0000	0,0000	0,0000
145	Samsung	Galaxy Tab 3 8,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
146	Samsung	Galaxy Tab 4 7,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
147	Samsung	Galaxy Tab 4 8,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
148	Samsung	Galaxy Tab 3 8,0	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
149	Samsung	Galaxy S4 Mini	<b>0,8123</b>	0,1877	0,0000	0,0000	0,0000

Tabel 5: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
150	Samsung	Galaxy S4 Zoom	0,0729	<b>0,9271</b>	0,0000	0,0000	0,0000
151	Samsung	Galaxy Mega 6,3	0,0518	<b>0,9469</b>	0,0000	0,0013	0,0000
152	Samsung	Galaxy Note 8	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
153	Samsung	Galaxy E7	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
154	Samsung	Galaxy S4 Active	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
155	Samsung	Galaxy A5	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
156	Samsung	Galaxy Note 3 Neo	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
157	Samsung	Galaxy Tab 4 10,1	0,0000	0,0000	0,0000	<b>1,0000</b>	0,0000
158	Samsung	Galaxy K Zoom	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
159	Samsung	Galaxy S4	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
160	Samsung	Galaxy Tab S 8,4	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
161	Samsung	Galaxy Note 2	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
162	Samsung	Galaxy S5	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
163	Samsung	Galaxy Tab S 10,5	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
164	Samsung	Galaxy Note 3	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
165	Samsung	Galaxy Alpha	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
166	Samsung	Galaxy Note 10,1	0,0000	0,0003	0,0000	<b>0,9997</b>	0,0000
167	Samsung	Galaxy Note 4	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
168	Samsung	Galaxy S6 Edge	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
169	Sony	Xperia E	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
170	Sony	Xperia Miro	0,0380	0,0000	<b>0,9620</b>	0,0000	0,0000
171	Sony	Xperia Tipo	0,0000	0,0000	<b>1,0000</b>	0,0000	0,0000
172	Sony	Xperia J	<b>0,5705</b>	0,0000	0,4295	0,0000	0,0000
173	Sony	Xperia E3	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
174	Sony	Xperia M	<b>0,9793</b>	0,0000	0,0207	0,0000	0,0000
175	Sony	Xperia Go	0,0816	0,0000	<b>0,9184</b>	0,0000	0,0000
176	Sony	Xperia L	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
177	Sony	Xperia M2	<b>0,9667</b>	0,0333	0,0000	0,0000	0,0000
178	Sony	Xperia C	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
179	Sony	Xperia P	<b>1,0000</b>	0,0000	0,0000	0,0000	0,0000
180	Sony	Xperia T2 Ultra	0,2652	<b>0,7348</b>	0,0000	0,0000	0,0000
181	Sony	Xperia M2 Aqua	<b>0,8750</b>	0,1250	0,0000	0,0000	0,0000
182	Sony	Xperia V	<b>0,9271</b>	0,0729	0,0000	0,0000	0,0000
183	Sony	Xperia T3	<b>0,8274</b>	0,1726	0,0000	0,0000	0,0000
184	Sony	Xperia C3	0,0037	<b>0,9963</b>	0,0000	0,0000	0,0000
185	Sony	Xperia ZR	0,0003	<b>0,9997</b>	0,0000	0,0000	0,0000
186	Sony	Xperia Z	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
187	Sony	Xperia Z Ultra	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000

Tabel 6: Nilai Taksiran Peluang Posterior Objek ( $\hat{h}(j|x_h)$ )

No	Merek	Tipe	1	2	3	4	5
188	Sony	Xperia ZL	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
189	Sony	Xperia Z1 Compact	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
190	Sony	Xperia Z1	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
191	Sony	Xperia Z3 Compact	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
192	Sony	Xperia Z2A	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
193	Sony	Xperia Z3 Tablet	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000
194	Sony	Xperia Z3	0,0000	<b>1,0000</b>	0,0000	0,0000	0,0000

## DAFTAR RIWAYAT HIDUP



**TEDY TRIYADI.** Lahir di Sumedang, 24 November 1992. Anak pertama dari pasangan Bapak Subiyanto dan Ibu Rohayati. Saat ini bertempat tinggal di Jalan Kayu Tinggi RT 006/ RW 12 No 28, Jakarta Timur 13910.

No. Ponsel : 082299790389

Email : tedytriyadiwijaya@gmail.com

**Riwayat Pendidikan :** Penulis mengawali pendidikan di SDI Al-Akhyar Jakarta pada tahun 1999 - 2005. Setelah itu, penulis melanjutkan ke SMP Negeri 193 Jakarta hingga tahun 2008. Kemudian kembali melanjutkan ke SMA Negeri 89 Jakarta dan lulus tahun 2011. Di Tahun yang sama penulis melanjutkan ke Universitas Negeri Jakarta (UNJ), jurusan Matematika, melalui jalur SNMPTN. Di pertengahan tahun 2015 penulis telah memperoleh gelar Sarjana Sains untuk Jurusan Matematika, Program Studi Matematika, FMIPA, UNJ.

**Riwayat Organisasi :** Di bangku sekolah, penulis aktif di berbagai organisasi, baik menjabat staff OSIS ketika SMP dan ketua OSIS ketika SMA, serta berbagai organisasi ilmiah. Selama di bangku perkuliahan, penulis aktif di berbagai organisasi kemahasiswaan. Dalam dua tahun pertama, penulis mendapat kepercayaan sebagai staff Departemen Akademik BEMJ Matematika, khususnya dalam pembuatan buku BASO++ (Bank Soal++), Seminar Skripsi, dan *Calculus Cup*.

**Riwayat Pekerjaan :** Penulis mulai menjadi pengajar matematika sejak tahun 2012. Pada tahun 2012, penulis menjadi pengajar matematika *private* dan beberapa bimbingan belajar, juga penulis pernah magang di PT. Pertamina Marketing and Budgeting Regional III Jawa Barat dan Bali sebagai *Strategic Sourcer*.

## SURAT PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya yang bertanda tangan di bawah ini, mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

Nama : Tedy Triyadi  
No. Registrasi : 3125111202  
Jurusan : Matematika  
Program Studi : Matematika

Menyatakan bahwa skripsi ini yang saya buat dengan judul "**Analisis Gerombol Kelas Laten (*Latent Class Cluster Analysis*) dalam Pengelompokan Beberapa Produk Ponsel Pintar *Smartphone***" adalah :

1. Dibuat dan diselesaikan oleh saya sendiri.
2. Bukan merupakan duplikat skripsi yang pernah dibuat oleh orang lain atau jiplakan karya tulis orang lain.

Pernyataan ini dibuat dengan sesungguhnya dan saya bersedia menanggung segala akibat yang timbul jika pernyataan saya tidak benar.

Jakarta, Juli 2015

Yang membuat pernyataan

Tedy Triyadi